Alexander Knop

Introduction to Discrete Mathematics

MARCH 28, 2019

Contents

List of Symbols

The letters A, B, X, Y, and Z denote sets, the letters x, y, and z denote the elements of X, Y, and Z respectively, P and Q denote propositions and predicates, the lower case latin letters f and g denote functions from X to Y and from Y to Z respectively, the letters a, b, n, and k denote integer numbers, and the greek letter α and β denote real numbers.

Counting

$(m)_n$	denotes the number of ways to choose an subset of n elements from a fixed set of n elements, page 75
$\binom{m}{n}$	denotes the number of ways to choose an unordered subset of n elements from a fixed set of m elements, page 76
$\lfloor \alpha \rfloor$	denotes the integer part of α, page 25
$\prod_{i=1}^{k} \alpha_i$	denotes $\alpha_1 \cdot \ldots \cdot \alpha_k$, page 27
$\sum_{i=1}^{k} \alpha_i$	denotes $\alpha_1 + \cdots + \alpha_k$, page 21
$\sum_{i \in S \,:\, P(i)} \alpha_i$	denotes $\alpha_{i_1} + \cdots + \alpha_{i_k}$, where $\{i \in S \;:\; P(i)\} = \{i_1, \ldots, i_k\}$, page 61
$n!$	denotes $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1$, page 27

Functions

$\mathrm{Im} f$	denotes the image of f, page 47
$f \circ g$	denotes the composition of funcitons f and g, page 46
$f\big\|_A$	denotes the restriction of f to the set A, page 45
f^{-1}	denotes the inverse of the funciton f (it's defined only when f is a bijection), page 59
$f^{-1}(y)$	depend on the context it either denotes the set $\{x \in X \;:\; f(x) = y\}$ if f is not a bijection and it denotes the value of f^{-1} at y if f is a bijection, page 59

I_A	denotes the identity function on the set A, page 46

Logical Notation

$\exists x \in X\ P(x)$	denotes the statement saying that P is true for some element of X, page 41
$\forall x \in X\ P(x)$	denotes the statement saying that P is true for all elements of X, page 41
$\neg P$	denotes the the statement saying that P is false, page 31
$P \implies Q$	denotes the statement saying that if P is true, then Q is true as well, page 9
$P \wedge Q$	denotes the the statement saying that P and Q are both true, page 31
$P \vee Q$	denotes the the statement saying that at least one of P and Q is true, page 30

Relations

$a \mid b$	says that a divides b, page 52
$a \equiv b \pmod{n}$	says that n divides $a - b$, page 50
$A \subseteq B$	says that A is a subset of B, page 34

Set Notation

$(B)_A$	denotes the set of injections from A to B, page 75
2^A	denotes the set of all the subsets of the set A, page 37
$[n]$	denotes the set of all the integers from 1 to n, page 35
$\bigcap_{i=1}^k A_i$	denotes $A_1 \cap \cdots \cap A_k$, page 38
$\bigcap_{i\in S\,:\,P(i)} A_i$	denotes $A_{i_1} \cap \cdots \cap A_{i_k}$, where $\{i \in S \,:\, P(i)\} = \{i_1, \ldots, i_k\}$, page 62
$\bigcup_{i=1}^k A_i$	denotes $A_1 \cup \cdots \cup A_k$, page 38
$\bigcup_{i\in S\,:\,P(i)} A_i$	denotes $A_{i_1} \cup \cdots \cup A_{i_k}$, where $\{i \in S \,:\, P(i)\} = \{i_1, \ldots, i_k\}$, page 62
$\binom{A}{k}$	denotes the set of subsets of A of cardinality k, page 76
$\mathbb{C}$	denotes the set of all complex numbers, page 33
$\emptyset$	denotes the set that does not have elements, page 34
$\mathbb{N}$	denotes the set of all integers greater than 0, page 33

$\mathbb{Q}$	denotes the set of all rational numbers, page 33
$\mathbb{R}$	denotes the set of all real numbers, page 33
$\mathbb{Z}$	denotes the set of all integers, page 33
$A \cap B$	denotes the intersection of two sets A and B, page 35
$A \cup B$	denotes the union of two sets A and B, page 35
$A \setminus B$	denotes the difference of two sets A and B, page 35
$A \times B$	denotes the set of all ordered pairs of elements of A and B, page 43
B^A	denotes the set of functions from A to B, page 75

Preface

- Why is a math book so sad?
- Because it’s full of problems.

Anonymous, Unknown

If you are reading this book, you probably have never studied proofs before. So let me give you some advice: mathematical books are very different from fiction, and even books in other sciences. Quite often you may see that some steps are missing, and some steps are not really explained and just claimed as obvious. The main reason behind this is to make the ideas of the proof more visible and to allow grasping the essence of proofs quickly.

Since the steps are skipped, you cannot just read the book and believe that you studied the topic; the best way to actually study the topic is to try to prove every statement before you read the actual proof in the book. In addition to this, I recommend trying to solve all the exercises in the book (you may find exercises in the middle and at the end of every chapter).

Additionally, many topics in this book have a corresponding five-minute video explaining the material of the chapter, it is useful to watch them before you go into the topic.

Organization

Part 1 covers the basics of mathematics and provide the language we use in the next parts. We start from the explanation of what a mathematical proof is (in Chapter 1). Chapter 2 shows how to prove theorems indirectly using proof by contradiction. Chapter 3 explains the most powerful method in our disposal, proof by induction. Finally, Chapters 4-7 define several important objects such as sets, functions, and relations.

Alexander Knop
San Diego, California, USA

Part I

Introduction to Mathematical Reasoning

1 Proofs

What is a Mathematical Proof:
Introduction to Mathematical Reasoning #1

youtu.be/eJD0gGqveIE

1.1 Direct Proofs

We start the discussion of the proofs in mathematics from an example of a proof in "everyday" life. Assume that we know that the following statements are true.

1. If a salmon has fins and scales it is kosher,
2. if a salmon has scales it has fins,
3. any salmon has scales.

Using these facts we may conclude that any salmon is kosher; indeed, any salmon has scales by the third statement, hence, by the second statement any salmon has fins, finally, by the first statement any salmon is kosher since it has fins and scales.

One may notice that this explanation is a sequence of conclusions such that each of them is true because the previous one is true. Mathematical proof is also a sequence of statements such that every statement is true if the previous statement is true. If P and Q are some statements and Q is always true when P is true, then we say that P implies Q. We denote the statement that P implies Q by $P \implies Q$.

In order to define the implication formally let us consider the following table.

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

Let P and Q be some statements. Then this table says that if P and Q are both false, then $P \implies Q$ is true etc.

Exercise 1.1. *Let n be an integer.*

1. *Is it always true that "n^2 is positive" implies "n is not equal to 0"?*

2. *Is it always true that "$n^2 - n - 2$ is equal to 0" implies "n is equal to 2"?*

In the example we gave at the beginning of the section we used some *known* facts. But what does it mean to know something? In math we typically say that we know a statement if we can prove it. But in order to prove this statement we need to know something again, which is a problem! In order to solve it, mathematicians introduced the notion of an *axiom*. An axiom is a statement that is believed to be true and when we prove a statement we prove it under the assumption that these axioms are true[1].

[1] Note that in different parts of math axioms may be different.

For example, we may consider axioms of inequalities for real numbers.

1. Let $a, b \in \mathbb{R}$. Only one of the following is true:
   - $a < b$,
   - $b < a$, or
   - $a = b$.
2. Let $a, b, c \in \mathbb{R}$. Then $a < b$ iff $a + c < b + c$ (iff is an abbreviation for "if and only if").
3. Let $a, b, c \in \mathbb{R}$. Then $a < b$ iff $ac < bc$ provided that $c > 0$ and $a < b$ iff $ac > bc$ if $c < 0$.
4. Let $a, b, c \in \mathbb{R}$. If $a < b$ and $b < c$, then $a < c$.

What We Know and How to Find a Proof: Introduction to Mathematical Reasoning #2

youtu.be/nBjJi6aTk2M

Let us now try to prove something using these axioms, we prove that if $a > 0$, then $a^2 > 0$. Note that $a > 0$, hence, by the third axiom $a^2 > 0$ (note that we also used an additionall statement saying that $0 \cdot 0 = 0$).

Similarly, we may prove that if $a < 0$, then $a^2 > 0$. And combining these two statements together we may prove that if $a \neq 0$, then $a^2 > 0$.

Such a way of constructing proof is called direct proofs.

Exercise 1.2. *Axiomatic system for a four-point geometry.*
Undefined terms: point, line, is on.
Axioms:

- *For every pair of distinct points x and y, there is a unique line ℓ such that x is on l and y is on l.*
- *Given a line ℓ and a point x that is not on ℓ, there is a unique line m such that x is on m and no point on ℓ is also on m.*
- *There are exactly four points.*
- *It is impossible for three points to be on the same line.*

Prove that there are at least two distinct lines.

Let n and m be some integers. Using direct proofs we may prove the following two statements.

- if n is even, then nm is also even (a number ℓ is even if there is an integer k such that $\ell = 2k$),
- if n is even and m is even, then $n + m$ is also even.

We start from proving the first statement. There is an integer k such that $n = 2k$ since n is even. As a result, $nm = 2(nk)$ so nm is even.

Now we prove the second statement. Since n and m are even there are k and ℓ such that $n = 2k$ and $m = 2\ell$. Hence, $n + m = 2(k + \ell)$ so $n + m$ is even.

1.2 *Constructing Proofs Backwards*

However, sometimes it is not easy to find the proof. In this case one of the possible methods to deal with this problem is to try to prove starting from the end.

For example, we may consider the statement $(a + b)^2 = a^2 + 2ba + b^2$. Imagine, for a second, that you have not learned about axioms. In this case you would write something like this:

$$\begin{aligned}(a+b)^2 = (a+b)\cdot(a+b) &= \\ a(a+b) + b(a+b) &= \\ a^2 + ab + ba + b^2 &= a^2 + 2ba + b^2.\end{aligned}$$

Let us try to prove it completely formally using the following axioms.

1. Let a, b, and c be reals. If $a = b$ and $b = c$, then $a = c$.
2. Let a, b, and c be reals. If $a = b$, then $a + c = b + c$ and $c + a = c + b$.
3. Let a, b, and c be reals. Then $a(b + c) = ab + ac$.
4. Let a and b be reals. Then $ab = ba$.
5. Let a and b be reals. Then $a + b = b + a$.
6. Let a be a real number. Then $a^2 = a \cdot a$ and $a \cdot a = a^2$.
7. Let a be a real number. Then $a + a = 2a$.

So the formal proof of the statement $(a + b)^2 = a^2 + 2ab + b^2$ is as follows. First note that $(a + b)^2 = (a + b) \cdot (a + b)$ (by axiom 6), hence, by axiom 1, it is enough to show that $(a + b) \cdot (a + b) = a^2 + 2ab + b^2$. By axiom 3, $(a + b) \cdot (a + b) = (a + b) \cdot a + (a + b) \cdot b$. Axiom 4 implies that $(a+b)\cdot a = a\cdot(a+b)$ and $(a+b)\cdot b = b\cdot(a+b)$ Hence, by axioms 1 and 2 applied twice

$$a\cdot(a+b)+b\cdot(a+b) = (a+b)\cdot a + b\cdot(a+b) = (a+b)\cdot a + (a+b)\cdot b.$$

As a result,

$$(a+b)\cdot(a+b) = (a+b)\cdot a + (a+b)\cdot b = \\ a\cdot(a+b)+b\cdot(a+b) = a\cdot a + a\cdot b + b\cdot a + b\cdot b;$$

so by axiom 1, it is enough to show that $a\cdot a + a\cdot b + b\cdot a + b\cdot b = a^2 + 2ab + b^2$. Additionally, by axiom 6, $a\cdot a = a^2$ and $b\cdot b = b^2$. Hence, by axiom 2, it is enough to show that $a^2 + a\cdot b + b\cdot a + b^2 = a^2 + 2ab + b^2$. By axiom 4, $a\cdot b = b\cdot a$, hence, by axiom 2, $a\cdot b + b\cdot a = b\cdot a + b\cdot a$. Therefore by axiom 7, $a\cdot b + b\cdot a = 2b\cdot a$. Finally, by axiom 2, $a\cdot b + b\cdot a + a^2 + b^2 = 2b\cdot a + a^2 + b^2$ and by axiom 5, $a\cdot b + b\cdot a + a^2 + b^2 = a^2 + a\cdot b + b\cdot a + b^2$ and $2b\cdot a + a^2 + b^2 = a^2 + 2b\cdot a + b^2$. Which finishes the proof by axiom 1.

1.3 *Analysis of Simple Algorithms*

We can use this knowledge to analyze simple algorithms. For example, let us consider the following algorithm. Let us prove that it is correct

Algorithm 1: The algorithm that finds the maximum element of a, b, c.

```
1: function Max(a, b, c)
2:     r ← a
3:     if b > r then
4:         r ← b
5:     end if
6:     if c > r then
7:         r ← c
8:     end if
9:     return r
10: end function
```

i.e. it returns the maximum of a, b, and c. We need to consider the following cases.

- If the maximum is equal to a. In this case, at line 2, we set $r = a$, at line 3 the inequality $b > r$ is false (since $a = r$ is the maximum) and at line 6 the inequality $c > r$ is also false (since $a = r$ is the maximum). Hence, we do not change the value of r after line 2 and the returned value is a.

- If the maximum is equal to b. We set $r = a$ at line 2. The inequality $b > r$ at line 3 is true (since b is the maximum) and we set r to be equal to b. So at line 6, the inequality $c > r$ is false (since $b = r$ is the maximum). Hence, the returned value is b.

- If the maximum is equal to c. We set $r = a$ at line 2. If the inequality $b > r$ is true at line 3 we set r to be equal to b. So at line 6 the inequality $c > r$ is true (since c is the maximum). Hence, we set r being equal to c and the returned value is c.

1.4 Proofs in Real-life Mathematics

In this chapter we explicitly used axioms to prove statements. However, it leads us to really long and hard to understand proofs (the last example in the previous section is a good example of this phenomenon). Because of this mathematicians tend to skip steps in the proofs when they believe that they are clear. It is worth to mention a nice quotation of Scott Aaronson about this problem

> When mathematicians say that a theorem has been "proved," they still mean, as they always have, something more like: "we've reached a social consensus that all the ideas are now in place for a strictly formal proof that could be verified by a machine ... with the only task remaining being massive rote coding work that none of us has any intention of ever doing!"

This is the reason why it is arduous to read mathematical texts and it is very different from reading non-mathematical books. A problem that arises because of this tendency is that some mistakes may happen if we skip way too many steps. In the last two centuries there were several attempts to solve this issue, one approach to this we are going to discuss in Part III.

Death of proof greatly exaggerated

`scottaaronson.com/blog/?p=4133`

End of The Chapter Exercises

1.3 Using the axioms of inequalities show that if a is a non-zero real number, then $a^2 > 0$.

1.4 Using the axioms of inequalities prove that for all real numbers a, b, and c,

$$bc + ac + ab \leq a^2 + b^2 + c^2.$$

1.5 Prove that for all integers a, b, and c, If a divides b and b divides c, then a divides c. Recall that an integer m divides an integer n if there is an integer k such that $mk = n$.

1.6 Show that square of an even integer is even.

1.7 Prove that 0 divides an integer a iff $a = 0$.

1.8 Using the axioms of inequalities, show that if $a > 0$, b, and c are real numbers, then $b \geq c$ implies that $ab \geq ac$.

1.9 Using the axioms of inequalities, show that if $a, b < 0$ are real numbers, then $a \leq b$ implies that $a^2 \geq b^2$.

2 Proofs by Contradiction

Proofs by Contradiction:
Introduction to Mathematical Reasoning #3

`youtu.be/bWP0VYx75DI`

2.1 Proving Negative Statements

The direct method is not very convenient when we need to prove a negation of some statement.

For example, we may try to prove that $78n + 102m = 11$ does not have integer solutions. It is not clear how to prove it directly since we can not consider all possible n and m. Hence, we need another approach. Let us assume that such a solution n, m exists. Note that $78n + 102m$ is even, but 11 is odd. In other words, an odd number is equal to an even number, it is impossible. Thus, the assumption was false.

Let us consider a more useful example, let us prove that if p^2 is even, then p is also even (p is an integer). Assume the opposite i.e. that p^2 is even but p is not. Let $p = 2b + 1$[1]. Note that $p^2 = (2b+1)^2 = 2(2b^2 + 2b) + 1$. Hence, p^2 is odd which contradicts to the assumption that p^2 is even.

[1] Note that we use here the statement that an integer n is not even iff it is odd, which, formally speaking, should be proven.

Using this idea we may prove much more complicated results e.g. one may show that $\sqrt{2}$ is irrational. For the sake of contradiction, let us assume that it is not true. In other words there are p and q such that $\sqrt{2} = \frac{p}{q}$ and $\frac{p}{q}$ is an irreducible fraction.

Note that $\sqrt{2}q = p$, so $2q^2 = p^2$. Which implies that p is even and 4 divides p^2. Therefore 4 divides $2q^2$ and q is also even. As a result, we get a contradiction with the assumption that $\frac{p}{q}$ is an irreducible fraction.

> **Template for proving a statement by contradiction.**
>
> Assume, for the sake of contradiction, that *the statement* is false. Then *present some argument that leads to a contradiction*. Hence, the assumption is false and *the statement* is true.

Exercise 2.1. *Show that* $\sqrt{3}$ *is irrational.*

2.2 Proving Implications by Contradiction

This method works especially well when we need to prove an implication. Since the implication $A \implies B$ is false only when A is true but B is false. Hence, you need to derive a contradiction from the fact that A is true and B is false.

We have already seen such examples in the previous section, we proved that p^2 is even implies p is even for any integer p. Let us consider another example. Let a and b be reals such that $a > b$. We need to show that $(ac < bc) \implies c < 0$. So we may assume that $ac < bc$ but $c \geq 0$. By the multiplicativity of the inequalities we know that if $(a > b)$ and $c > 0$, then $ac > bc$ which contradicts to $ac < bc$.

A special case of such a proof is when we need to prove the implication $A \implies B$, assume that B is false and derive that A is false which contradicts to A (such proofs are called proofs by contraposition); note that the previous proof is a proof of this form.

2.3 Proof of "OR" Statements

Another important case is when we need to prove that at least one of two statements is true. For example, let us prove that $ab = 0$ iff $a = 0$ or $b = 0$. We start from the implication from the right to the left. Since if $a = 0$, then $ab = 0$ and the same is true for $b = 0$ this implication is obvious.

The second part of the proof is the proof by contradiction. Assume $ab = 0$, $a \neq 0$, and $b \neq 0$. Note that $b = \frac{ab}{a} = 0$, hence $b = 0$ which is a contradiction to the assumption.

End of The Chapter Exercises

2.2 Prove that if n^2 is odd, then n is odd.

2.3 In Euclidean (standard) geometry, prove: If two lines share a common perpendicular, then the lines are parallel.

2.4 Let us consider four-lines geometry, it is a theory with undefined terms: point, line, is on, and axioms:

1. there exist exactly four lines,
2. any two distinct lines have exactly one point on both of them, and
3. each point is on exactly two lines.

Show that every line has exactly three points on it.

2.5 Let us consider group theory, it is a theory with undefined terms: group-element and times (if a and b are group elements, we denote a times b by $a \cdot b$), and axioms:

1. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for every group-elements a, b, and c;
2. there is a unique group-element e such that $e \cdot a = a = a \cdot e$ for every group-element a (we say that such an element is the identity element);
3. for every group-element a there is a group-element b such that $a \cdot b = e$, where e is the identity element;
4. for every group-element a there is a group-element b such that $b \cdot a = e$, where e is the identity element.

Let e be the identity element. Show the following statements

- if $b_0 \cdot a = b_1 \cdot a = e$, then $b_0 = b_1$, for every group-elements a, b_0, and b_1.
- if $a \cdot b_0 = a \cdot b_1 = e$, then $b_0 = b_1$, for every group-elements a, b_0, and b_1.
- if $a \cdot b_0 = b_1 \cdot a = e$, then $b_0 = b_1$, for every group-elements a, b_0, and b_1.

2.6 Let us consider three-points geometry, it is a theory with undefined terms: point, line, is on, and axioms:

1. There exist exactly three points.
2. Two distinct points are on exactly one line.
3. Not all the three points are collinear i.e. they do not lay on the same line.
4. Two distinct lines are on at least one point i.e. there is at least one point such that it is on both lines.

Show that there are exactly three lines.

2.7 Show that there are irrational numbers a and b such that a^b is rational.

2.8 Show that there does not exist the largest integer.

2.9 Let us consider Young's geometry, it is a theory with undefined terms: point, line, is on, and axioms:

1. there exists at least one line,
2. every line has exactly three points on it,
3. not all points are on the same line,

4. for two distinct points, there exists exactly one line on both of them,
5. if a point does not lie on a given line, then there exists exactly one line on that point that does not intersect the given line.

Show that for every point, there are exactly four lines on that point.

3 Proofs by Induction

The Induction Principle:
Introduction to Mathematical Reasoning #4

youtu.be/jOnZTWGpX_I

3.1 Simple Induction

Let us consider a simple problem: what is bigger 2^n or n? In this chapter, we are going to study the simplest way to prove that $2^n > n$ for all positive integers n. First, let us check that it is true for small integers n.

	1	2	3	4	5	6	7	8
n	1	2	3	4	5	6	7	8
2^n	2	4	8	16	32	64	128	256

We may also note that 2^n is growing faster than n, so we expect that if $2^n > n$ for small integers n, then it is true for all positive integers n.

In order to prove this statement formally, we use the following principle.

Principle 3.1 (The Induction Principle)**.** *Let $P(n)$ be some statement about a positive integer n. Hence, $P(n)$ is true for every positive integer n iff*

(the base case) $P(1)$ *is true and*

(the induction step) $P(k) \implies P(k+1)$ *is true for all positive integers k.*

Let us prove now the statement using this principle. We define $P(n)$ be the statement that "$2^n > n$". $P(1)$ is true since $2^1 > 1$. Let us assume now that $2^n > n$. Note that $2^{n+1} = 2 \cdot 2^n > 2n \geq n+1$. Hence, we proved the induction step.

Exercise 3.1. *Prove that $(1+x)^n \geq 1 + nx$ for all positive integers n and real numbers $x \geq -1$.*

3.2 Changing the Base Case

Let us consider functions n^2 and 2^n.

	1	2	3	4	5	6	7	8
n^2	1	4	9	16	25	36	49	64
2^n	2	4	8	16	32	64	128	256

Note that 2^n is greater than n^2 starting from 5. But without some trick we can not prove this using induction since for $n = 3$ it is not true!

The trick is to use the statement $P(n)$ stating that $(n+4)^2 < 2^{n+4}$. The base case when $n = 1$ is true. Let us now prove the induction step. Assume that $P(k)$ is true i.e. $(k+4)^2 < 2^{k+4}$. Note that $2(k+4)^2 < 2^{k+1+4}$ but $(k+5)^2 = k^2 + 10k + 25 \leq 2k^2 + 16k + 32 = 2(k+4)^2$. Which implies that $2^{k+1+4} > (k+5)^2$. So $P(k+1)$ is also true.

In order to avoid this strange $+4$ we may change the base case and use the following argument.

Theorem 3.1. *Let $P(n)$ be some statement about an integer n. Hence, $P(n)$ is true for every integer $n > n_0$ iff*

(the base case) $P(n_0 + 1)$ *is true and*

(the induction step) $P(k) \implies P(k+1)$ *is true for all integers $k > n_0$.*

Using this generalized induction principle we may prove that $2^n \geq n^2$ for $n \geq 4$. The base case for $n = 4$ is true. The induction step is also true; indeed let $P(k)$ be true i.e. $(k+4)^2 < 2^{k+4}$. Hence, $2(k+4)^2 < 2^{k+1+4}$ but $(k+5)^2 = k^2 + 10k + 25 \leq 2k^2 + 16k + 32 = 2(k+4)^2$.

Let us now prove the theorem. Note that the proof is based on an idea similar to the trick with $+4$, we just used.

Proof of Theorem 3.1. $\Rightarrow$ If $P(n)$ is true for any $n > n_0$ it is also true for $n = n_0 + 1$ which implies the base case. Additionally, it true for $n = k + 1$ so the induction step is also true.

$\Leftarrow$ In this direction the proof is a bit harder. Let us consider a statement $Q(n)$ saying that $P(n + n_0)$ is true. Note that by the base case for P, $Q(1)$ is true; by the induction step for P we know that $Q(n)$ implies $P(n+1)$. As a result, by the induction principle $Q(n)$ is true for all positive integers n. Which implies that $P(n)$ is true for all integers $n > n_0$.

$\square$

3.3 *Inductive Definitions*

We may also define objects inductively. Let us consider the sum $1 + 2 + \cdots + n$ a line of dots indicating "and so on" which indicates the definition by induction. In this case, a more precise notation is $\sum_{i=1}^{n} i$.

Definition 3.1. *Let $a(1), \ldots, a(n), \ldots$ be a sequence of integers. Then $\sum_{i=1}^{n} a(i)$ is defined inductively by the following statements:*

- $\sum_{i=1}^{1} a(i) = a(1)$, *and*
- $\sum_{i=1}^{k+1} a(i) = \sum_{i=1}^{k} a(i) + a(k+1)$.

Let us prove that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Note that by definition $\sum_{i=1}^{1} i = 1$ and $\frac{1(1+1)}{2} = 1$; hence, the base case holds. Assume that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Note that $\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + (n+1)$ and by the induction hypothesis $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Hence, $\sum_{i=1}^{n+1} i = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$.

Exercise 3.2. *Prove that* $\sum_{i=1}^{n} 2^i = 2^{n+1} - 2$*.*

3.4 *Analysis of Algorithms with Cycles*

Induction is very useful for analysing algorithms using cycles. Let us extend the example we considered in Section 1.3.

Let us consider the following algorithm. We prove that it is working

Algorithm 2: The algorithm that finds the maximum element of $a_1, \ldots, a_n$.

```
function MAX(a_1, ..., a_n)
    r ← a_1
    for i from 2 to n do
        if a_i > r then
            r ← a_i
        end if
    end for
    return r
end function
```

correctly. First, we need to define $r_1, \ldots, r_n$ the value of r during the execution of the algorithm. It is easy to see that $r_1 = a_1$ and

$$r_{i+1} = \begin{cases} r_i & \text{if } r_i > a_{i+1} \\ a_{i+1} & \text{otherwise} \end{cases}.$$

Secondly, we prove by induction that r_i is the maximum of $a_1, \ldots, a_i$. It is clear that the base case for $i = 1$ is true. Let us prove the induction step from k to $k+1$. By the induction hypothesis, r_k is the maximum of $a_1, \ldots, a_k$. We may consider two following cases.

- If $r_k > a_{k+1}$, then $r_{k+1} = r_k$ is the maximum of $a_1, \ldots, a_{k+1}$ since r_k is the maximum of $a_1, \ldots, a_k$.

- Otherwise, a_{k+1} is greater than or equal to $a_1, \ldots, a_k$, hence, $r_{k+1} = a_{k+1}$.

Exercise 3.3. *Show that line 6 in the following sorting algorithm executes* $\frac{n(n+1)}{2}$ *times.*

Algorithm 3: The algorithm is selection sort, it sorts $a_1, \ldots, a_n$.

```
function SelectionSort(a_1, ..., a_n)
    for i from 1 to n do
        r ← a_i
        ℓ ← i
        for j from i to n do
            if a_j > r then
                r ← a_j
                ℓ ← j
            end if
        end for
        Swap a_i and a_ℓ.
    end for
end function
```

3.5 *Strong Induction*

Sometimes $P(k)$ is not enough to prove $P(k+1)$ and we need all the statements $P(1), \ldots, P(k)$. In this case we may use the following induction principle.

Theorem 3.2 (The Strong Induction Principle). *Let $P(n)$ be some statement about positive integer n. Hence, $P(n)$ is true for every integer $n > n_0$ iff*

(the base case) $P(n_0+1)$ *is true and*

(the induction step) If $P(n_0+1), \ldots, P(n_0+k)$ are true, then $P(n_0+k+1)$ is also true for all positive integers k.

Before we prove this theorem let us prove some properties of Fibonacci numbers using this theorem. The Fibonacci numbers are defined as follows: $f_0 = 0$, $f_1 = 1$, and $f_k = f_{k-1} + f_{k-2}$ for $k \geq 2$ (note that they are also defined using strong induction since we use not only f_{k-1} to define f_k).

Theorem 3.3 (The Binet formula). *The Fibonacci numbers are given by the following formula*

$$f_n = \frac{\alpha^n - \beta^n}{\sqrt{5}},$$

where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

Proof. We use the strong induction principle to prove this statement with $n_0 = -1$. Let us first prove the base case, $\frac{(\alpha^0-\beta^0)}{\sqrt{5}} = 0 = f_0$. We also need to prove the induction step.

- If $k = 1$, then $\frac{(\alpha^1-\beta^1)}{\sqrt{5}} = 1 = f_1$.

- Otherwise, by the induction hypothesis, $f_k = \frac{\alpha^k - \beta^k}{\sqrt{5}}$ and $f_{k-1} = \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}$. By the definition of the Fibonacci numbers $f_{k+1} = f_k + f_{k-1}$. Hence,
$$f_{k+1} = \frac{\alpha^k - \beta^k}{\sqrt{5}} + \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}.$$
Note that it is enough to show that
$$\frac{\alpha^{k+1} - \beta^{k+1}}{\sqrt{5}} = \frac{\alpha^k - \beta^k}{\sqrt{5}} + \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}. \tag{3.1}$$
Note that it is the same as
$$\frac{\alpha^{k+1} - \alpha^k - \alpha^{k-1}}{\sqrt{5}} = \frac{\beta^{k+1} - \beta^k - \beta^{k-1}}{\sqrt{5}}.$$
Additionally, note that α and β are roots of the equation $x^2 - x - 1 = 0$. Hence, $\alpha^{k+1} - \alpha^k - \alpha^{k-1} = \alpha^{k-1}(\alpha^2 - \alpha - 1) = 0$ and $\beta^{k+1} - \beta^k - \beta^{k-1} = \beta^{k-1}(\beta^2 - \beta - 1) = 0$. Which implies equality (3.1).

□

Now we are ready to prove the strong induction principle.

Proof of Theorem 3.2. It is easy to see that if $P(n)$ is true for all $n > n_0$, then the base case and the induction steps are true. Let us prove that if the base case and the induction step are true, then $P(n)$ is true for all $n > n_0$.

Let $Q(k)$ be the statement that $P(n_0 + 1), \ldots, P(n_0 + k)$ are true. Note that $Q(1)$ is true by the base case for P. Additionally, note that if $Q(k)$ is true, then $Q(k+1)$ is also true, by the induction step for P. Hence, by the induction principle, $Q(k)$ is true for all positive integers k. Which implies that $P(n_0 + k)$ is true for all positive integers k. □

3.6 *Recursive Definitions*

Sometimes you wish to define objects using objects of the same form like in the case of inductive definitions but you do not know how to enumerate them using an integer parameter.

One example of such a situation is the definition of an arithmetic formula.

(the base case) x_i is an arithmetic formula on the variables $x_1, \ldots, x_n$ for all i; if c is a real number, then c is also an arithmetic formula on the variables $x_1, \ldots, x_n$.

recursion step: If P and Q are arithmetic formulas on the variables $x_1, \ldots, x_n$, then $(P + Q)$ and $P \cdot Q$ are arithmetic formulas on the variables $x_1, \ldots, x_n$.

Note that this definition implicitly states that any other expressions are not arithmetic formulas.

We can define recursively the value of such a formula. Let $v_1, \dots, v_n$ be some integers.

(the base cases) $x_i\big|_{x_1=v_1,\dots,x_n=v_n} = v_i$; in other words, the value of the arithmetic formula x_i is equal to v_i when $x_1 = v_1, \dots, x_n = v_n$; if c is a real number, then $c|_{x_1=v_1,\dots,x_n=v_n} = c$.

(the recursion steps) If P and Q are arithmetic formulas on the variables $x_1, \dots, x_n$, then

$$(P+Q)\big|_{x_1=v_1,\dots,x_n=v_n} = P\big|_{x_1=v_1,\dots,x_n=v_n} + Q\big|_{x_1=v_1,\dots,x_n=v_n}$$

and

$$(P \cdot Q)\big|_{x_1=v_1,\dots,x_n=v_n} = P\big|_{x_1=v_1,\dots,x_n=v_n} \cdot Q\big|_{x_1=v_1,\dots,x_n=v_n}.$$

For example, $((x_1 + x_2) \cdot x_3)$ is clearly an arithmetic formula on the variables $x_1, \dots, x_n$. One may expect the value of this formula with $x_1 = 1$, $x_2 = 0$, and $x_3 = -1$ be equal to -1, let us check:

- Note that
$$\begin{aligned} x_1\big|_{x_1=1,x_2=0,x_3=-1} &= 1, \\ x_2\big|_{x_1=1,x_2=0,x_3=-1} &= 0, \text{ and} \\ x_3\big|_{x_1=1,x_2=0,x_3=-1} &= -1. \end{aligned}$$
- Hence,
$$(x_1 + x_2)\big|_{x_1=1,x_2=0,x_3=-1} = 1 + 0 = 1.$$
- Finally,
$$((x_1 + x_2) \cdot x_3)\big|_{x_1=1,x_2=0,x_3=-1} = 1 \cdot -1 = -1.$$

A special case of induction which called structural induction is the easiest way to prove properties of recursively defined objects. The idea of this is similar to the idea of strong induction:

- first, we prove the statement for the base case,
- after that we prove the induction step, using the assumption that the statement is true for all the substructures (e.g. subformulas in the previous definition).

To illustrate this method, we prove the following theorem.

Theorem 3.4. *For any arithmetic formula A on x, there is a polynomial p such that $p(v) = A\big|_{x=v}$ for any real value v.*

Proof. (the base cases) If $A = x_i$, then consider the polynomial $p(x) = x$; it is easy to see that $A\big|_{x=v} = v = p(v)$. If $A = c$ where c is a real number, then consider the constant polynomial $p(x) = c$; it is easy to note that $A\big|_{x=v} = c = p(v)$.

(the induction step) We need to consider two cases. Consider the case when $A = B_1 + B_2$. By the induction hypothesis, there are polynomials q_1 and q_2 such that $B_1\big|_{x=v} = q_1(v)$ and $B_2\big|_{x=v} = q_2(v)$ for all real numbers v. We define $p(x) = q_1(x) + q_2(x)$ (it is a polynomial since sum of two polynomials is a polynomial). It is obvious that $A\big|_{x=v} = B_1\big|_{x=v} + B_2\big|_{x=v} = q_1(v) + q_2(v) = p(v)$.

Another case is $A = B_1 \cdot B_2$. Again, by the induction hypothesis, there are polynomials q_1 and q_2 such that $B_1\big|_{x=v} = q_1(v)$ and $B_2\big|_{x=v} = q_2(v)$ for all real numbers v. We define $p(x) = q_1(x) \cdot q_2(x)$ (it is a polynomial since product of two polynomials is a polynomial). It is obvious that $A\big|_{x=v} = B_1\big|_{x=v} \cdot B_2\big|_{x=v} = q_1(v) \cdot q_2(v) = p(v)$.

□

Exercise 3.4. • *Define arithmetic formulas with division and define their value (make sure that you handled divisions by* 0*).*

• *Show that for any arithmetic formula with division A on x, there are polynomials p and q such that $\frac{p(v)}{q(v)} = A\big|_{x=v}$ or $A\big|_{x=v}$ is not defined for any real value v.*

3.7 *Analysis of Recursive Algorithms*

To illustrate the power of recursive definitions and strong induction, let us analyze Algorithm 4. We prove that number of comparisons of this algorithm is bounded by $6 + 2\log_2(n)$. First step of the proof is to denote the worst number of comparisons when we run the algorithm on the list of length n by $C(n)$. It is easy to see that $C(n) = n$ for $n \le 5$. Additionally, $C(n) \le 1 + \max(C(\lfloor \frac{n}{2} \rfloor), C(n - \lfloor \frac{n}{2} \rfloor))$ for $n > 5$. As we mentioned we prove that $C(n) \le 6 + 2\log_2(n)$, we prove it by induction. The base case is clear; let us now prove the induction step. By the induction hypothesis,

$$C(\left\lfloor \frac{n}{2} \right\rfloor) \le 6 + 2\log_2(\left\lfloor \frac{n}{2} \right\rfloor)$$

and

$$C(n - \left\lfloor \frac{n}{2} \right\rfloor) \le 6 + 2\log_2(n - \left\lfloor \frac{n}{2} \right\rfloor),$$

where $\lfloor \alpha \rfloor$ denotes the integer part of a real number α. Since $\lfloor \frac{n}{2} \rfloor \le \frac{n}{2}$ and $n - \lfloor \frac{n}{2} \rfloor \le \frac{n}{2} + 1$, $C(n) \le 1 + 2\log_2(\frac{n}{2} + 1)$. However,

$$1 + 6 + 2\log_2\left(\frac{n}{2} + 1\right) \le 6 + 2\log_2\left(\frac{n}{\sqrt{2}} + \sqrt{2}\right) \le 6 + 2\log_2(n)$$

```
function BinarySearch(e, a_1, ..., a_n)
    if n ≤ 5 then
        for i from 1 to n do
            if a_i = e then
                return i
            end if
        end for
    else
        ℓ ← ⌊n/2⌋
        if a_ℓ ≤ e then
            BinarySearch(e, a_1, ..., a_ℓ)
        else
            BinarySearch(e, a_{ℓ+1}, ..., a_n)
        end if
    end if
end function
```

Algorithm 4: The binary search algorithm that finds an element e in the sorted list $a_1, \ldots, a_n$.

for $n \geq 5$. As a result, we proved the induction step.

End of The Chapter Exercises

3.5 Show that there does not exist the largest integer.

3.6 Show that for any positive integer n, $n^2 + n$ is even.

3.7 Show that for any positive integer n, 3 divises $n^3 + 2n$.

3.8 Show that for any integer $n \geq 10$, $n^3 \leq 2^n$.

3.9 Show that for any positive integer n, $\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x}$.

3.10 Show that for any matrix $A \in \mathbb{R}^{m \times n}$ $(n > m)$ there is a nonzero vector $x \in \mathbb{R}^n$ such that $Ax = 0$.

3.11 Show that all the elements of $\{0,1\}^n$ (Binary strings) may be ordered such that every successive strings in this order are different only in one character. (For example, for $n = 2$ the order may be 00, 01, 11, 10.)

3.12 Let $a_0 = 2$, $a_1 = 5$, and $a_n = 5a_{n-1} - 6a_{n-2}$ for all integers $n \geq 2$. Show that $a_n = 3^n + 2^n$ for all integers $n \geq 0$.

3.13 Show that $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$ for all integers $n \geq 1$.

3.14 Show that $\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$ for all integers $n \geq 1$.

3.15 Show that $\sum_{i=1}^{n} \frac{1}{i^2} \leq 2 - \frac{1}{n}$ for all integers $n \geq 1$.

3.16 Show that $\sum_{i=1}^{n}(2i-1) = n^2$ for any positive integer n.

3.17 Prove that $\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$ for any positive integer n.

3.18 Prove that $\sum_{i=2}^{n}(i+1)2^i = n2^{n+1}$ for all integers $n > 2$.

3.19 Let $a_1, \ldots, a_n$ be a sequence of real numbers. We define inductively $\prod_{i=k}^{n} a_i$ as follows:

- $\prod_{i=1}^{1} a_i = a_1$ and
- $\prod_{i=1}^{k+1} a_i = \left(\prod_{i=1}^{k} a_i\right) \cdot a_{k+1}$.

Prove that $\prod_{i=1}^{n-1}\left(1 - \frac{1}{(i+1)^2}\right) = \frac{n+1}{2n}$ for all integers $n > 1$.

3.20 Let $f_0 = 1$, $f_1 = 1$, and $f_{n+2} = f_{n+1} + f_n$ for all integers $n \geq 0$. Show that $f_n \geq \left(\frac{3}{2}\right)^{n-2}$.

3.21 Show that $f_{n+m} = f_{n-1}f_{m-1} + f_n f_m$.

3.22 Show that two arithmetic formulas $(x_1 + x_2) \cdot x_3$ and $x_1 \cdot x_3 + x_2 \cdot x_3$ on the variables x_1, x_2, and x_3 have the same values.

3.23 We say that L is a list of powers of x iff

- either $L = x^k$ for some positive integer k or
- $L = (x^k, L')$ where L' is a list of powers of x and k is a positive integer.

Let L be a list of powers of x. We say that the sum of L with $x = v$ denoted by $\sum L\big|_{x=v}$

- is equal to x^k whether $L = x^k$ and
- is equal to $x^k + \sum L'\big|_{x=v}$ whether $L = (x^k, L')$.

Prove that for any list L of powers of x there is a polynomial such that $\sum L\big|_{x=v} = p(v)$ for all real numbers v.

3.24 Let us define $n!$ as follows: $1! = 1$ and $n! = (n-1)! \cdot n$. Show that $n! \geq 2^n$ for any $n \geq 4$.

3.25 Show that $\int\limits_{0}^{+\infty} x^n e^{-x}\, \mathrm{d}x = n!$ for all $n \geq 0$.

3.26 Prove that $\sum_{i=1}^{n}(i+1)2^i = n2^{n+1}$ for all integers $n \geq 1$.

3.27 Show that $\sum_{k=1}^{n} k \cdot k! = (n+1)! - 1$.

4 Predicates and Connectives

Connectives and Propositions:
Introduction to Mathematical Reasoning #5

youtu.be/0unvlq2OTaE

4.1 Propositions and Predicates

In the previous chapters we used the word "statement" without any even relatively formal definition of what it means. In this chapter we are going to give a semi-formal definition and discuss how to create complicated statements from simple statements.

It is difficult to give a formal definition of what a mathematical statement is, hence, we are not going to do it in this book. The goal of this section is to enable the reader to recognize mathematical statements.

A *proposition* or a mathematical statement is a declarative sentence which is either true or false but not both. Consider the following list of sentences.

1. $2 \times 2 = 4$
2. $\pi = 4$
3. n is even
4. 32 is special
5. The square of any odd number is odd.
6. The sum of any even number and one is prime.

Of those, the first two are propositions; note that this says nothing about whether they are true or not. Actually, the first is true and the second is false. However, the third sentence becomes a proposition only when the value of n is fixed. The fourth is not a proposition. Finally, the last two are propositions (the fifth is true and the sixth is false).

The third statement is somewhat special, because there is a simple way to make it a proposition: one just needs to fix the value of the variables. Such sentences are called predicates and the variables that need to be specified are called free variables of these predicates.

Note that the fourth sentence is also interesting, since if we define what it means to be special, the phrase became a proposition. Mathematicians tend to do such things to give mathematical meanings to everyday words.

4.2 *Connectives*

Mathematicians often need to decide whether a given proposition is true or false. Many statements are complicated and constructed from simpler statements using *logical connectives*. For example we may consider the following statements:

1. $3 > 4$ and $1 < 1$;

2. $1 \times 2 = 5$ or $6 > 1$.

Logical connective "OR". The second statement is an example of usage of this connective. The statement "P or Q" is true if and only if at least one of P and Q is true. We may define the connective using the truth table of it.

P	Q	P or Q
T	T	T
T	F	T
F	T	T
F	F	F

The or connective is also called *disjunction* and the disjunction of P and Q is often dented as $P \vee Q$.

Warning: Note that in everyday speech "or" is often used in the exclusive case, like in the sentence "we need to decide whether it is an insect or a spider". In this case the precise meaning of "or" is made clear by the context. However, mathematical language should be formal, hence, we always use "or" inclusively.

Logical connective "AND". The first statement is an example of this connective. The statement "P and Q" is true if and only if both P and Q are true. We may define the connective using the truth table of it.

P	Q	P and Q
T	T	T
T	F	F
F	T	F
F	F	F

The or connective is also called *conjunction* and the conjunction of P and Q is often dented as $P \wedge Q$.

> **Warning:** Not all the properties of "and" from everyday speach are captured by logical conjunction. For example, "and" sometimes implies order. For example, "They got married and had a child" in common language means that the marriage came before the child. The word "and" can also imply a partition of a thing into parts, as "The American flag is red, white, and blue." Here it is not meant that the flag is at once red, white, and blue, but rather that it has a part of each color.

Logical connective "NOT". The last connective is called *negation* and examples of usage of it are the following:

1. 5 is not greater than 8;
2. Does not exist an integer n such that $n^2 = 2$.

Note that it is not straightforward where to put the negation in these sentences.

The negation of a statement P is denoted as $\neg P$ (sometimes it is also denoted as $\sim P$).

End of The Chapter Exercises

4.1 Construct truth tables for the statements

- not (P and Q);
- (not P) or (not Q);
- P and (not Q);
- (not P) or Q;

4.2 Consider the statement "All gnomes like cookies". Which of the following statements is the negation of the above statement?

- All gnomes hate cookies.
- All gnomes do not like cookies.
- Some gnome do not like cookies.
- Some gnome hate cookies.
- All creatures who like cookies are gnomes.
- All creatures who do not like cookies are not gnomes.

4.3 Using truth tables show that the following statements are equivalent:

- $P \implies Q$,
- $(P \vee Q) \iff Q$ ($A \iff B$ is the same as $(A \implies B) \wedge (B \implies A)$),
- $(P \wedge Q) \iff P$

4.4 Prove that three connectives "or", "and", and "not" can all be written in terms of the single connective "notand" where "P notand Q" is interpreted as "not (P and Q)" (this operation is also known as Sheffer stroke or NAND).

4.5 Show the same statement about the connective "notor" where "P notor Q" is interpreted as "not (P or Q)" (this operation is also known as Peirce's arrow or NOR).

5 Sets

Sets:
Introduction to Mathematical Reasoning #6

youtu.be/bshBV2H4Sqo

5.1 The Intuitive Definition of a Set

A set is one of the two most important concepts in mathematics. Many mathematical statements involve "an integer n" or "a real number a". Set theory notation provides a simple way to express that a is a real number. However, this language is much more expressible and it is impossible to imagine modern mathematics without this notation.

As in the previous chapter it is difficult to define a set formally so we give a less formal definition which should be enough to use the notation. A *set* is a well-defined collection of objects. Important examples of sets are:

1. $\mathbb{R}$ a set of reals,
2. $\mathbb{Z}$ the set of integers[1],
3. $\mathbb{N}$ the set of natural numbers[2],
4. $\mathbb{Q}$ a set of rational numbers,
5. $\mathbb{C}$ a set of complex numbers.

[1] "Z" stands for the German word Zahlen ("numbers").

[2] Note that in the literature there are two diffirent traditions: in one 0 is a natural number, in another it is not; in this book we are going to assume that 0 is not a natural number.

Usually, sets are denoted by single letter.

Objects in a set are called *elements* of the set and we denote the statement "x is in the set E" by the formula $x \in E$ and the negation of this statement by $x \notin E$. For example, we proved that $\sqrt{2} \notin \mathbb{Q}$[3].

[3] The symbol $\in$ was first used by Giuseppe Peano 1889 in his work "Arithmetices principia, nova methodo exposita". Here he wrote on page X: "The symbol $\in$ means is. So $a \in b$ is read as a *is* a b; ..." The symbol itself is a stylized lowercase Greek letter epsilon ("ϵ"), the first letter of the word εστι, which means "is".

Exercise 5.1. *Which of the following sets are included in which? Recall that a number is prime iff it is an integer greater than* 1 *and divisible only by* 1 *and itself.*

1. *The set of all positive integers less than* 10.
2. *The set of all prime numbers less than* 11.
3. *The set of all odd numbers greater than* 1 *and less than* 6.
4. *The set of all positive integers less than* 10.
5. *The set whose only elements are* 1 *and* 2.

6. *The set whose only element is* 1.

7. *The set of all prime numbers less than* 11.

5.2 Basic Relations Between Sets

Many problems in mathematics are problems of determining whether two description of sets are describing the same set or not. For example, when we learn how to solve quadratic equations of the form $ax^2 + bx + c = 0$ ($a, b, c \in \mathbb{R}$) we learn how to list the elements of the set $\{x \in \mathbb{R} : ax^2 + bx + c = 0\}$.

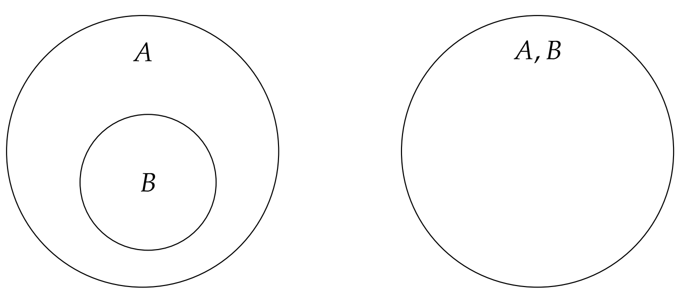
We say that two sets A and B are equal if they contain the same elements (we denote it by $A = B$). If all the elements of A belong to B we say that A is a subset of B and denote it by $A \subseteq B$[4].

For example, $\mathbb{Q} \subseteq \mathbb{R}$ since any rational number is also a real number. A special set is an empty set i.e. the set that does not have elements, we denote it $\emptyset$.

[4] In the literature there are three symbols for "subset": $\subseteq$, $\subsetneq$, and $\subseteq$. $A \subseteq B$ means that A is a subset of B and we allow $A = B$ and $A \subsetneq B$ means that A is a subset of B and we forbid $A = B$. However, there is a problem with the third symbol, some people use it as a synonym of $\subseteq$ and some use it as a synonym of $\subsetneq$. Due to this ambiguity we are going to avoid using it in this book.

Diagrams

If we think of a set A as represented by all the points within a circle or any other closed figure, then it is easy to represent the notion of A being a subset of another set B also represented by all the points within a circle. We just put a circle labeled by A inside of the circle labeled by B. We can also diagram an equality by drawing a circle labeled by both A and B. (see fig. 5.1). Such diagrams are called Euler diagrams and it is clear that one may draw Euler diagrams for more than two sets.



(a) $A \subseteq B$ (b) $A = B$

Figure 5.1: Euler diagrams for subset and equality relations

Descriptions of Sets

In this section we describe how to define new sets, this notation is also known as *set-builder notation*.

Listing elements. The simplest way to define a set is just to list the elements. For example

1. $\{1, 2, \pi\}$ is the set consisting of three elements 1, 2, and π, and
2. $\{1, 2, 3, \ldots\}$ is the set of all positive integers i.e. it is the set $\mathbb{N}$.

Conditional definitions. We may also describe a set using some constraint e.g we may list all the even numbers using the following formula $\{n \in \mathbb{Z} \ : \ n \text{ is even}\}$ (we read it as "the set of all integers n such that n is even").

Using this we may also define the set of all integers from 1 to m, we denote it $[m]$; i.e. $[m] = \{n \in \mathbb{N} \ : \ 0 < n \leq m\}$.

Constructive definitions. Another way to construct a set of all even numbers is to use the constructive definition of a set: $\{2k \ : \ k \in \mathbb{Z}\}$.

We may also describe a set of rational numbers using this description: $\mathbb{Q} = \{a/b \ : \ a \in \mathbb{Z}, b \in \mathbb{N}\}$ (note that we may also use a mix of a conditional and constructive definitions, $\mathbb{Q} = \{a/b \ : \ a, b \in \mathbb{Z}, b \neq 0\}$).

Exercise 5.2. *Describe a set of perfect squares using constructive type of definition.*

Disjoint Sets

Two sets are *disjoint* iff they do not have common elements. We also say that two sets are *overlapping* iff they are not disjoint i.e. they share an element.

More generally, $A_1, \ldots, A_\ell$ are pairwise disjoint iff A_i is disjoint with A_j for all $i \neq j \in [\ell]$

Exercise 5.3. *Of the sets in Exercise 5.1, which are disjoint from which?*

5.3 *Operations over Sets.*

Another way to describe a set is to apply operation to other sets. Let A and B be sets.

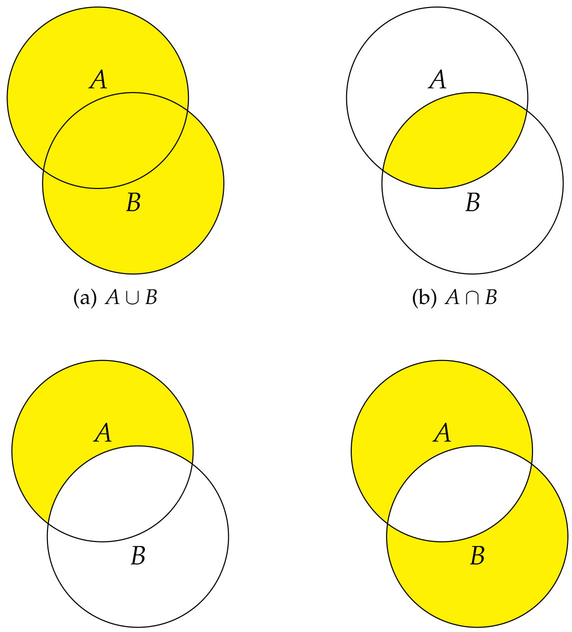
The first example of the operations on sets is the *union* operation. The union of A and B is the set containing all the elements of A and all the elements of B i.e. $A \cup B = \{x \ : \ x \in A \text{ or } x \in B\}$[5].

Another example of such an operation is *intersection*. The intersection of A and B is the set of all the elements belonging to both A and B i.e $A \cap B = \{x \ : \ x \in A \text{ and } x \in B\}$[6].

The third operation we are going to discuss this lecture is *set difference*. If A and B are some sets, then $A \setminus B = \{x \ : \ x \in A \text{ and } x \notin B\}$.

[5] Note that this definition is not correct since in the conditional definitions we have to specify the set x belongs to and we cannot do this here.

[6] You may notice that in the definition of the union we use disjunction and in the definition of intersection we use conjunction. Actually this is a the reason the symbol of the conjunction is similar to the symbol of intersection and the symbol of the disjunction is similar to the symbol of union.



(a) $A \cup B$ (b) $A \cap B$

(c) $A \setminus B$ (d) $A \Delta B$

Figure 5.2: Euler diagrams for set operations

The last operation is *symmetric difference*. If A and B are some sets, then $A\Delta B = (A \setminus B) \cup (B \setminus A)$. Note that alternatively $A\Delta B = (A \cup B) \setminus (A \cap B)$

Exercise 5.4. *Describe the set* $\{n \in \mathbb{N} \ : \ n \text{ is even}\} \cap \{3n \ : \ n \in \mathbb{N}\}$.

Theorem 5.1. *Let A, B, and C be some sets. Then we have the following identities.*

associativity: $A \cup (B \cup C) = (A \cup B) \cup C$ *and* $A \cap (B \cap C) = (A \cap B) \cap C$.

commutativity: $A \cup B = B \cup A$ *and* $A \cap B = B \cap A$.

distributivity: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *and* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. One may prove these properties using the Euler diagrams. Alternatively they can be proven by definitions. Let us prove only the first part of the distributivity, the rest is Exercise 5.5.

Our proof consists of two parts in the first part we prove that $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$. Suppose that $x \in A \cup (B \cap C)$. Then $x \in A$ or $x \in (B \cap C)$.

- If $x \in A$, then $x \in (A \cup B)$ and $x \in (A \cup C)$ i.e. $x \in ((A \cup B) \cap (A \cup C))$.
- If $x \in (B \cap C)$, then $x \in B$ and $x \in C$. Which implies that $x \in (A \cup B)$ and $x \in (A \cup C)$. As a result, $x \in ((A \cup B) \cap (A \cup C))$.

□

Exercise 5.5. *Prove the rest of the equalities in Theorem 5.1.*

Probably the most difficult concept connected to sets is the concept of a power set. Let A be some set, then the set of all possible subsets of A is denoted by 2^A (sometimes this set is denoted by $\mathcal{P}(A)$) and called the power set of A. In other words $2^A = \{B \, : \, B \subseteq A\}$.

> **Warning:** Please do not forget about two extremal elements of the power set 2^A: the empty set and A itself.

For example if $A = \{1, 2, 3\}$, then

$$2^A = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

5.4 The Well-ordering Principle

Using the set notation we may finally justify the proof of the statement that $2^n > n$ for all positive integers n from the video about mathematical induction. In order to do this let us first formulate the following theorem.

Theorem 5.2. *Let $A \subseteq \mathbb{Z}$ be a non-empty set. We say that $b \in \mathbb{Z}$ is a lower bound for the set A iff $b \leq a$ for all $a \in A$. Additionally, we say that the set A is bounded if there is a lower bound for A.*

Given this, if A is bounded, then there is a lower bound $a \in A$ for the set A (we say that a is the minimum of the set A).

Note that this theorem also states that any subset of natural numbers have a minimum.

Recall that we wish to prove that $2^n > n$ for all positive n. Assume that it is not true, in this case the set $A = \{n \in \mathbb{N} \, : \, 2^n < n\}$ is non-empty. Denote by n_0 the minimum of the set A, n_0 exists by Theorem 5.2. We may consider the following two cases.

- If $n_0 = 1$, then it leads to a contradiction since $2 = 2^1 > 1$.
- Otherwise, note that $1 \leq n_0 - 1 < n_0$, hence, $2^{n_0 - 1} > n_0 - 1$. So $2^{n_0} > 2n_0 - 2 \geq n_0$. Which is a contradiction with the definition of n_0.

Finally, we prove Theorem 5.2.

Proof of Theorem 5.2. Let b be a lower bound for the set A. Assume that there is no minimum of the set A. Let $P(n)$ be the statement that $n \notin A$.

First, we are going to prove that $P(n)$ is true for all $n \geq b$. The base case is true since if $b \in A$, then b is the minimum of A which contradicts to the assumption that there is no minimum of A. The induction step is also clear, by the induction hypothesis we know that $P(b)$, $\ldots$, $P(k)$ are true, hence, $(k+1) \in A$ implies that $k+1$ is the minimum of A.

Now we prove that A is empty. Assume the opposite i.e. assume that there is $x \in A$. Note that $x \geq b$ since b is a lower bound of A. However, $P(x)$ is true which implies that $x \notin A$. Therefore the assumption was false and A is empty, but this contradicts to the fact that A is non- empty. $\square$

End of The Chapter Exercises

5.6 Find the power sets of $\emptyset$, $\{1\}$, $\{1,2\}$, $\{1,2,3,4\}$. How many elements in each of this sets?

5.7 Prove that

- $A \subseteq B \iff A \cup B = B$,
- $A \subseteq B \iff A \cap B = A$.

5.8 Let A be a subset of a set U we call this set a universe. We say that the set $\overline{A} = U \setminus A$ is a complement of A in U. Show the following equalities

- $\overline{\overline{A}} = A$.
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$.
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

5.9 Let us define an intersection of more than two sets as follows. Let $A_1, \ldots, A_n$ be some sets. Then

- $\bigcap_{i=1}^{1} A_i = A_1$ and
- $\bigcap_{i=1}^{k+1} A_i = \left(\bigcap_{i=1}^{k} A_i\right) \cap A_{k+1}$.

Show that $\bigcap_{i=1}^{n} \{x \in \mathbb{N} \,:\, i \leq x \leq n\} = \{n\}$ for all integers $n > 0$.

5.10 Let us define a union of more than two sets as follows. Let A_1, $\ldots$, A_n be some sets. Then

- $\bigcup_{i=1}^{1} A_i = A_1$ and
- $\bigcup_{i=1}^{k+1} A_i = \left(\bigcup_{i=1}^{k} A_i\right) \cup A_{k+1}$.

Show that $\bigcup_{i=1}^{n} [i] = [n]$ for all integers $n > 0$.

5.11 Let Ω be some set and $A_1, \ldots, A_n \subseteq \Omega$. Show that $\bigcup_{i=1}^{n} A_i = \{x \in \Omega \,:\, \exists i \in [n]\; x \in A_i\}$.

5.12 Let $A_1, \dots, A_n$ be some sets. Show that $\bigcup_{i=1}^{n}(A_i \cap B) = (\bigcup_{i=1}^{n} A_i) \cap B$.

5.13 Show that $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

6 Functions

Functions and Quantifiers:
Introduction to Mathematical Reasoning #7

youtu.be/VHJeUrCedTU

Another important type of objects in mathematics are functions. Function f from a set X to a set Y (written as $f : X \to Y$) is a unique assignment of elements of Y to the elements of X (note that it is not necessary that all the elements of Y are used). In other words, for each element $x \in X$ there is one assigned element $f(x) \in Y$. We call such an element the *value* of f at x, we also say that $f(x)$ is an *image* of x.

Unfortunately, the definition is not formal. Through this chapter we will provide a more formal definition.

6.1 *Quantifiers.*

The first ingredient is called quantifiers. Very often we use phrases like "all the people in the class have smartphones." However, we still do not know how to write it using symbols.

The Universal Quantifier. In order to say "all" or "every" we use the symbol $\forall$[1]: if $P(a)$ is a predicate about $a \in A$, then $\forall a \in A\ P(a)$ is a statement saying that all the elements of A satisfy the predicate P. In other words it is the same as the statement $\{a \in A : P(a)\} = A$. For example, $\forall x \in \mathbb{R}\ x \cdot 0 = 0$ says that product of every real number and zero is equal to zero.

[1] The symbol is a turned "A" symbol, the first letter of the word "all".

The Existential Quantifier. The second quantifier means "there is" and is denoted by the symbol $\exists$[2]: if $P(a)$ is a predicate about an element of A, then $\exists a \in A\ P(a)$ says that there is an element of A satisfying the predicate P i.e. $\{a \in A : P(a)\} \neq \emptyset$. For example, $\exists x \in \mathbb{R}\ x^2 - 1 = 0$ states that there is a real solution of the equation $x^2 - 1 = 0$.

[2] The symbol is a turned "E" symbol, the first letter of the word "exists". It is also intresting that the symbol for the universal quantifier was introduced by Gerhard Gentzen in 1935 but the symbol for the existencial qunatifier was intriduced, 38 years earlier, by Giuseppe Peano in 1897.

Warning: Note that the word "any" sometimes indicates a universal statement and sometimes an existential statement.
Standard meaning of "any" is "every" like in the statement "$a^2 \geq 0$ for any real number", therefore this statement can be rewritten as $\forall a \in \mathbb{R}\ a^2 \geq 0$. Nonetheless, in the negative and interrogative statements "any" is used to mean "some". For example, "There is not any real number a such that $a^2 < 0$" is asserting that the statement $\exists a \in \mathbb{R}\ a^2 < 0$ is false. And "Is there any real number a such that $a^2 = 1$?" is asking whether the existential statement $\exists a \in \mathbb{R}\ a^2 = 1$ is true.
Real care is required with questions involving "any": "Is there any integer a such that $a \geq 1$?" clearly is asking whether $\exists a \in \mathbb{R}\ a^2 \geq 1$ is true; however, "Is $a \geq 1$ for any integer a" is less clear and might be taken to asking about the same question as the first question, $\exists a \in \mathbb{Z}\ a \geq 1$ (which is true) but might also be taken to be asking about $\forall a \in \mathbb{Z}\ a \geq 1$ (which is false).

Proving Statements Involving Quantifiers

Most of the statements in mathematics involve quantifiers. This is one of the factors distinguishing advanced from elementary mathematics. In this section we give an overview of the main methods of proof. Though the whole book is about proving such results.

Proving statements of the form $\forall a \in A\ P(a)$. Such statements can be rewritten in the form $a \in A \implies P(a)$. For example, we proved earlier that $a^2 \geq 0$ for all real numbers a using this approach.

Proving statements of the form $\exists a \in A\ P(a)$. The easiest way to prove such a statement is by simply exhibiting an element a of A such that $P(a)$ is true. This method is called *proof by example*.

Let us prove the statement $\exists x \in \mathbb{N}\ x^2 = 4$ using this method. Observe that $2 \in N$ and $2^2 = 4$ so $x = 2$ provides an example proving this statement. There are, however, less direct methods such as use of the counting arguments.

Proving statements involving both quantifiers. To illustrate problems of this type let us prove that for any integer n, if n is even, then n^2 is also even.

This statement is a universal statement $\forall n \in \mathbb{Z}\ (n \text{ is even} \implies n^2 \text{ is even})$. However, the hypothesis that n is even is an existential statement $\exists q \in \mathbb{Z}\ n = 2q$. So we begin the proof as follows:

Suppose that n is an even integer. Then $n = 2q$ for some integer q.

The conclusion we wish to prove is that n^2 is even, which may be written as $\exists q \in \mathbb{Z}\ n^2 = 2q$. Note that q here is a dummy variable used to express the statement n^2 is a doubled integer. We may replace it by any other letter not already in use, for example $\exists p \in \mathbb{Z}\ n^2 = 2p$. Hence, if we present p such that $n^2 = 2p$, we finish the proof. As a result, we can complete the proof as follows.

Therefore, $n^2 = (2q)^2 = 4q^2$ and so, since $2q^2$ is an integer n^2 is even.

Disproving Statements Involving Quantifiers

Disproving something seems a bit off from the first glance, but to some extent it is the same as proving the negation.

Disproving statements of the form $\forall a \in A\ P(a)$. We may note that the negation of such a statement is the statement $\exists a \in A\ \neg P(a)$. So we can disprove it by giving a single example for which it is false. This is called *disproof by counterexample* to $P(a)$.

For example, we may disprove the statement $\forall x \in \mathbb{R}\ x^2 > 2$ by giving a counterexample $x = 1$ since $1^2 = 1 < 2$.

Disproving statements of the form $\exists a \in A\ P(a)$. The negation of this statement is the statement $\forall a \in A\ \neg P(a)$. Which gives one way of disproving the statement.

Let us prove that does not exist a real number x such that $x^2 = -1$. We know that, for all $x \in \mathbb{R}$, we have the inequality $x^2 \geq 0$ and so $x^2 \neq -1$. Hence, there does not exist $x \in \mathbb{R}$ such that $x^2 = -1$.

6.2 *Cartesian product*

Another ingredient is the notion of Cartesian product. If X and Y are two sets, then $X \times Y = \{(x, y)\ :\ x \in X \text{ and } y \in Y\}$. We also denote $\underbrace{X \times X \times \cdots \times X}_{k \text{ times}}$ by X^k.

Consider the following example. If $X = \{a, b, c\}$ and $Y = \{a, b\}$, then

$$X \times Y = \{(a, a), (a, b), (b, a), (b, b), (c, a), (c, b)\}.$$

Additionally, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is the familiar 2-dimensional Euclidean plane.

Exercise 6.1. *Find the set $\{a, b\} \times \{a, b\} \setminus \{(x, x)\ :\ x \in \{a, b\}\}$*

Theorem 6.1. *For all sets A, B, C, and D the following hold:*

- $A \times (B \cup C) = (A \times B) \cup (A \times C)$;
- $A \times (B \cap C) = (A \times B) \cap (A \times C)$;

- $(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D)$;
- $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.

Proof. It is easy to prove this statement by the definitions. Let us prove only the second equality, the rest is Exercise 6.2.

Note that $(x, y) \in A \times (B \cap C)$ iff $x \in A$ and $y \in (B \cap C)$. Hence, $(x, y) \in A \times (B \cap C)$ iff $x \in A$, $y \in B$, and $y \in C$. Thus $(x, y) \in A \times (B \cap C)$ iff $(x, y) \in (A \times B)$ and $(x, y) \in (A \times C)$. As a result, $(x, y) \in A \times (B \cap C)$ iff $(x, y) \in (A \times B) \cap (A \times C)$ as required. □

Exercise 6.2. *Prove the rest of the equalities in Theorem 6.1.*

6.3 *Graphs of Functions*

Now we have all the components to define a function. Mathematicians think about the functions in the way we defined them at the beginning of the chapter, however formally in order to define a function $f : X \to Y$ one need to define a set $D \subseteq X \times Y$ (such a set is called the *graph of the function* f) such that

- $\forall x \in X\ \exists y \in Y\ (x, y) \in D$ and
- $\forall x \in X, y_1, y_2 \in Y\ ((x, y_1) \in D \wedge (x, y_2) \in D \implies y_1 = y_2)$.

We say that $y \in Y$ is the value $f(x)$ of the function described by D at $x \in X$ iff $(x, y) \in D$.

The simplest way to think about the functions is in the terms of tables. Let us use this idea to list all the functions $\{a, b, c\}$ to $\{d, e\}$.

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$f_6(x)$	$f_7(x)$	$f_8(x)$
a	d	d	d	d	e	e	e	e
b	d	d	e	e	d	d	e	e
c	d	e	d	e	d	e	d	e

Exercise 6.3. *List all the functions from* $\{a, b\}$ *to* $\{a, b\}$.

However, listing all the values of a function is only possible when the domain of the function is finite. Thus the most common way to describe a function is using a formula which provides a way to find the value of a function. When the function is defined as a formula it is important to be clear which sets are the domain and the codomain of the function.

Let $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Consider the following functions.

- $g_1 : \mathbb{R} \to \mathbb{R}$ such that $g_1(x) = x^2$;
- $g_2 : \mathbb{R}_+ \to \mathbb{R}$ such that $g_2(x) = x^2$;

- $g_3 : \mathbb{R} \to \mathbb{R}_+$ such that $g_3(x) = x^2$;
- $g_4 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $g_4(x) = x^2$;

Nonetheless that all these functions are defined using the same formula x^2, we will see in the next chapters that these four functions have different properties.

Exercise 6.4. *Find the graph of the function $f : \mathbb{Z} \to \mathbb{Z}$ such that $f(x) = 3x$.*

Note that when you define the function you need to define it such that the definition makes sense for all the elements of the domain. For example, the formula $g(x) = \frac{x^2-3x+2}{x-1}$ does not define a function from $\mathbb{R}$ to $\mathbb{R}$ since it is not defined for $x = 1$. It is typical to define a function from real numbers to real numbers by a formula and the convention is that the domain is the set of all numbers for which the formula makes sense (unless the domain is specified explicitly). Using this convention the formula g defines a function from $\mathbb{R} \setminus \{1\}$ to $\mathbb{R}$.

If we really need a function from $\mathbb{R}$ there are two possible approaches for extending g.

Rewriting the formula. We can rewrite the formula such that it makes sense for all the real numbers. Note that for all $x \in \mathbb{R} \setminus \{1\}$,

$$\frac{x^2 - 3x + 2}{x - 1} = \frac{(x-2)(x-1)}{x-1} = x - 2.$$

Then $g_1(x) = x - 2$ defines a function on $\mathbb{R}$ extending the function g.

Explicit definition. Alternatively we can explicitly specify the value of g at 1. So

$$g_2(x) = \begin{cases} \frac{x^2-3x+2}{x-1} & \text{if } x \neq 1 \\ -1 & \text{if } x = 1 \end{cases}$$

defines a function from $\mathbb{R}$ to $\mathbb{R}$. Note that we can specify the values at individual points any way we want.

Similarly to sets we may define the equality between functions. We say that two functions $f, g : X \to Y$ are equal ($f = g$) iff $f(x) = g(x)$ for all $x \in X$ i.e. their graphs are equal. Note that two functions are equal only if they have the same domains and codomains. For example, g_1 and g_2 we just defined are equal to each other none the less that we defined them in two different ways.

We defined g_1 and g_2 to extend g to a bigger domain, similarly we can make a domain smaller.

Definition 6.1. *Let $f : X \to Y$ and $A \subseteq X$. Then $f|_A : A \to Y$ is a function such that $\forall x \in A\ f|_A(x) = f(x)$ (we say that $f|_A$ is the restriction of f to the set A).*

6.4 Composition of Functions

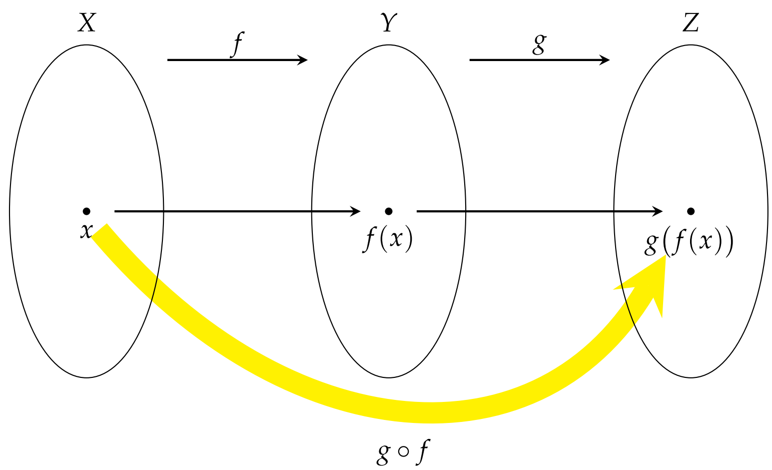


Figure 6.1: Composition of functions

Suppose $f : X \to Y$ and $g : Y \to Z$ be some function. Then, given an element $x \in X$, the function f assigns $y = f(x) \in Y$, and the function g assigns $z = g(y) = g(f(x)) \in Z$. Thus using f and g an element of Z can be assigned to x. This operation defines a function from X to Z and the result of this operation is called the *composition* of f and g.

Definition 6.2. *If $f : X \to Y$ and $g : Y \to Z$, then $h = g \circ f$ is a function from X to Z such that $g(f(x)) = h(x)$ for all $x \in X$.*

Let us consider an example. Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x + 1$ and $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = x^2$. Then $(g \circ f) : \mathbb{R} \to \mathbb{R}$ and $(g \circ f)(x) = (x+1)^2$ for all $x \in \mathbb{R}$. Note that the order of f and g is important since $(f \circ g)(x) = x^2 + 1$. Thus composition is not *commutative*.

There are two special type functions.

- Let $A \subseteq X$, then $i : A \to X$ such that $i(a) = a$ for all $a \in A$ is called the *inclusion* function of A into X. Observe that $(f \circ i) : A \to Y$ and $(f \circ i) = f|_A$ for any function $f : X \to Y$.

- Another important function is called the *identity* function. Let X be some set. Then $I_X : X \to X$ is the identity function on X iff $I_X(x) = x$.

Theorem 6.2. *Let $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$. Then*

- $h \circ (g \circ f) = (h \circ g) \circ f$.

- $f \circ I_X = f = I_Y \circ f$.

Proof. These results can be proven simply by evaluating the functions. For example, both functions in the first equality assign $h(g(f(x)))$ for any $x \in X$ and so functions are equal. $\square$

Notice that this theorem states that we may write $f \circ g \circ h$ without ambiguity.

6.5 The Image of a Function

Given a function $f : X \to Y$, it is not necessary that every element of Y is an image of some $x \in X$. For example, the function $\mathbb{R} \to \mathbb{R}$ defined by the formula x^2 does not have -1 as a value.

Thus we may give the following definition.

Definition 6.3. *The image of the function f is defined as follows*

$$\mathrm{Im} f = \{y \in Y \,:\, \exists x \in X\; f(x) = y\} = \{f(x) \,:\, x \in X\}$$

(in other words it is the projection of the graph D of f on the second coordinate: $\mathrm{Im} f = \{y \,:\, (x, y) \in D\}$*).*

End of The Chapter Exercises

6.5 Find an image of the function $f : \mathbb{Z} \to \mathbb{Z}$ such that $f(x) = 3x$.

6.6 Determine the following sets:

- $\{m \in \mathbb{N} \,:\, \exists n \in \mathbb{N}\; m \le n\}$;
- $\{m \in \mathbb{N} \,:\, \forall n \in \mathbb{N}\; m \le n\}$;
- $\{n \in \mathbb{N} \,:\, \exists m \in \mathbb{N}\; m \le n\}$;
- $\{n \in \mathbb{N} \,:\, \forall m \in \mathbb{N}\; m \le n\}$.

6.7 Prove or disprove the following statements.

- $\forall m, n \in \mathbb{N}\; m \le n$.
- $\exists m, n \in \mathbb{N}\; m \le n$.
- $\exists m \in \mathbb{N} \forall n \in \mathbb{N}\; m \le n$.
- $\forall m \in \mathbb{N} \exists n \in \mathbb{N}\; m \le n$.
- $\exists n \in \mathbb{N} \forall m \in \mathbb{N}\; m \le n$.
- $\forall n \in \mathbb{N} \exists m \in \mathbb{N}\; m \le n$.

7 *Relations*

Nonetheless that function are used almost everywhere in mathematics, many relations are not functional by their nature. For example, could never define a function $r(a)$ that gives the solution of $x^2 = a$ because there are two solutions for $a > 0$ and there are zero solutions for $a < 0$. A relation is a more general mathematical object.

In order to define a relation we need to relax the definition of the graph of a function (Section 6.3) by allowing more than one "result" and by allowing zero "results". In other words we just say that any set $R \subseteq X_1 \times \cdots \times X_k$ is a k-ary *relation* on $X_1, \ldots, X_k$. We also say that $x_1 \in X_1, \ldots, x_k \in X_k$ are in the relation R iff $(x_1, \ldots, x_k) \in R$. If $k = 2$ such a relation is called a *binary relation* and we write xRy if x and y are in the relation R. If $X_1 = \cdots = X_k = X$, we say that R is a k-ary relation on X.

Note that $=$, $\leq$, $\geq$, $<$, and $>$ define relations on $\mathbb{R}$ (or any subset S of $\mathbb{R}$). For example, if $S = \{0, 1, 2\}$, then $<$ defines the relation $R = \{(0, 1), (0, 2), (1, 2)\}$.

Probably the most popular relation in mathematics is the following relation on $\mathbb{Z}$. Let $a, b \in \mathbb{Z}$. If n divides $a - b$ for some $n \in Z$, we say that "a equivalent to b modulo n" and denote it as $a \equiv b \pmod{n}$. For example, 1 and 4 are equivalent modulo 3 since 3 divides $1 - 4 = -3$.

7.1 *Equivalence Relations*

The definition of a relation is way to broad. Hence, quite often we consider some types of relation. Probably the most interesting type of the relations is equivalence relations.

Definition 7.1. *Let R be a relation on a set X. We say that R is an* equivalence relation *if it satisfies the following conditions:*

reflexivity: xRx for any $x \in X$;

symmetry: xRy iff yRx for any $x, y \in X$;

transitivity: for any $x, y, z \in X$, if xRy and yRZ, then xRz;

One may guess that the equivalence relation are mimicking $=$, so it is not a surprise that $=$ is an equivalence relation.

The definition seems quite bizarre, however, all of you are already familiar with an important example: you know that equivalent fractions represent the same number. For example $\frac{2}{4}$ is the same as $\frac{1}{2}$. Let us consider this example more thorough, let S be a set of symbols of the form $\frac{x}{y}$ (note that it is not a set of numbers) where $x, y \in Z$ and $y \neq 0$. We define a binary relation R on S such that $\frac{x}{y}$ and $\frac{z}{w}$ are in the relation R iff $xw = zy$. It is easy to prove that this relation is an equivalence relation.

reflexivity: Let $\frac{a}{b} \in S$. Since $ab = ab$, we have that $\frac{a}{b} R \frac{a}{b}$.

symmetry: Let $\frac{a}{b}, \frac{c}{d} \in S$. Suppose that $\frac{a}{b} R \frac{c}{d}$, by the definition of R, it implies that $ac = db$. As a result, $\frac{c}{d} R \frac{a}{b}$.

transitivity: Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in S$ with $\frac{a}{b} R \frac{c}{d}$ and $\frac{c}{d} R \frac{e}{f}$. Then $ad = cb$ and $cf = ed$. The first equality can be rewritten as $c = ad/b$. Hence, $adf/b = ed$ and $af = eb$ since $d \neq 0$. So $\frac{a}{b} R \frac{e}{f}$.

Partitions

Let S be some set. We say that $\{P_1, \ldots, P_k\}$ form a partition of S iff P_1, $\ldots$, P_k are pairwise disjoint and $P_1 \cup \cdots \cup P_k = S$; in other words, a partition is a way of dividing a set into overlapping pieces.

Exercise 7.1. *Let $\{P_1, \ldots, P_k\}$ be a partition of a set S and R be a binary relation of S such that aRb iff $a, b \in P_i$ for some $i \in [k]$. Show that R is an equivalence relation.*

This exercise shows that one may transform a partition of the set S into an equivalence relation on S. However, it is possible to do the opposite.

Theorem 7.1. *Let R be a binary equivalence relation on a set S. For any element $x \in S$, define $R_x = \{y \in S : xRy\}$ (the set of all the elements of S related to x) we call such a set the equivalence class of x. Then $\{R_x : x \in S\}$ is a partition of S.*

Exercise 7.2. *Prove Theorem 7.1.*

Modular Arithmetic

The relation "$\equiv \pmod{n}$" is actively used in the number theory. One of the important properties of this relation is that it is an equivalence relation.

Theorem 7.2. *The relation $\equiv \pmod{n}$ is an equivalence relation.*

Proof. To prove this statement we need to prove all three properties: reflexivity, symmetry, and transitivity.

reflexivity: Note that for any integer x, $x - x = 0$ is divisible by any integer including n. Hence, $x \equiv x \pmod{n}$.

symmetry: Let us assume that $x \equiv y \pmod{n}$; i.e. $x - y = kn$ for some integer k. Note that $y - x = (-k)n$, so $y \equiv x \pmod{n}$.

transitivity: finally, assume that $x \equiv y \pmod{n}$ and $y \equiv z \pmod{n}$; i.e. $x - y = kn$ and $y - z = \ell n$ for some integers k and ℓ. It is easy to note that $x - z = (x - y) + (y - z) = (k + \ell)n$. As a result, $x \equiv z \pmod{n}$.

Thus, we proved that $\equiv \pmod{n}$ is an equivalence relation. □

Let $x \in \mathbb{Z}$; we denote by $r_{x,n}$ the equivalence class of x with respect to the relation $\equiv \pmod{n}$, we also denote by $\mathbb{Z}/n\mathbb{Z}$ the set of all the equivalence classes with respect to the relation $\equiv \pmod{n}$.

Another important property of these relation is that they behave well with respect to the arithmetic operations.

Theorem 7.3. *Let $x, y \in Z$ and $n \in \mathbb{N}$. Suppose that $a \in r_{x,n}$ and $b \in r_{y,n}$, then $(a + b) \in r_{x+y,n}$ and $ab \in r_{xy,n}$.*

Using this theorem we may define arithmetic operations on the equivalence classes with respect to the relation $\equiv \pmod{n}$. Let $x, y \in \mathbb{Z}$ and $n \in \mathbb{N}$. Then $r_{x,n} + r_{y,n} = \{a + b \ : \ a \in r_{x,n}, b \in r_{y,n}\} = r_{x+y,n}$ and $r_{x,n}r_{y,n} = \{ab \ : \ a \in r_{x,n}, b \in r_{y,n}\} = r_{xy,n}$. Moreover, these operations have plenty of good properties.

Exercise 7.3. *Let $a, b, c \in \mathbb{Z}/n\mathbb{Z}$. Show that the following equalities are true:*

- $a + (b + c) = (a + b) + c$,
- $a + r_{0,n} = a$ *(thus we denote $r_{0,n}$ as 0),*
- $ar_{1,n} = a$ *(thus we denote $r_{1,n}$ as 1),*
- *there is a class $d \in \mathbb{Z}/n\mathbb{Z}$ such that $a + d = r_{0,n}$ (thus we denote this d as $-a$),*
- $a + b = b + a$,
- $ab = ba$,
- $a(b + c) = ab + ac$,

7.2 *Partial Orderings*

In the previous section we discussed a mathematical way to express the property being similar. In this section we are going to give a way to analyze relation similar to comparisons.

Definition 7.2. *A binary relation R on S is a partial ordering if it satisfies the following constraints.*

reflexivity: xRx *for any* $x \in S$*;*

antisymmetry: *if* xRy *and* yRx*, then* $x = y$ *for all* $x, y \in S$*;*

transitivity: *for any* $x, y, z \in S$*, if* xRy *and* yRZ*, then* xRz*;*

We say that an order R on a set S is total iff for any $x, y \in S$, either xRy or yRx.

Note that if S is a set of numbers, then $\leq$ defines a partial ordering on S; moreover, it defines a total order.

Typically we use symbols similar to $\preceq$ to denote partial orderings and we write $a \prec b$ to express that $a \preceq b$ and $a \neq b$.

Let $|$ be the relation on $\mathbb{Z}$ such that $d \mid n$ iff d divides n.

Theorem 7.4. *The relation $|$ is a partial ordering of the set $\mathbb{N}$.*

Proof. To prove that this relation is a partial ordering we need to check all three properties.

reflexivity: Note that $x = 1 \cdot x$ for any integer x; hence, $x \mid x$ for any integer x.

antisymmetry: Assume that $x \mid y$ and $y \mid x$. Note that it means that $kx = y$ and $\ell y = x$ for some integers k and ℓ. Hence, $y = (k \cdot \ell)y$ which implies that $k \cdot \ell = 1$ and $k = \ell = 1$. Thus, $x = y$.

transitivity: finally, assume that $x \mid y$ and $y \mid z$; i.e. $kx = y$ and $\ell y = z$. As a result, $(k \cdot \ell)x = z$ and $x \mid z$.

$\square$

Exercise 7.4. *Let S be some set, show that $\subseteq$ defines a partial ordering on the set 2^S.*

Topological Sorting

Partial orderings are very useful for describing complex processes. Suppose that some process consists of several tasks, T denotes the set of these tasks. Some tasks can be done only after some others e.g. when you cooking a salat you need to wash vegetables before you chop them. If $x, y \in T$ be some tasks, $x \preceq y$ if x should be done before y and this is a partial ordering.

In the applications this order is not a total order because some steps do not depend on other steps beeing done first (you can chop tomatos and chop cucumbers in any order). However, if we need to create a schedule in which the tasks should be done, we need to create a total ordering on T. Moreover, this order should be compatible with the partial ordering. In other words, if $x \preceq y$, then $x \preceq_t y$ for all $x, y \in T$, where $\preceq_t$ is the total order. The technique of finding such a total ordering is called *topological sorting*.

Theorem 7.5. *Let S be a finite set and $\preceq$ be a partial order on S. Then there is a total order $\preceq_t$ on S such that if $x \preceq y$, then $x \preceq_t y$ for all $x, y \in S$*

This sorting can be done using the following procedure.

- Initiate the set S beeing equal to T
- Choose the minimal element of the set S with respect to the ordering $\preceq$ (such an element exists since S is a finite set, see Chapter 8). Add this element to the list, remove it from the set S, and repeate this step if $S \neq \emptyset$.

Let us consider the following example. In the left column we list the classes and in the right column the prerequisite.

Courses	Prerequisite
Math 20A	
Math 20B	Math 20A
Math 20C	Math 20B
Math 18	
Math 109	Math 20C, Math 18
Math 184A	Math 109

We need to find an order to take the courses.

1. We start with

   $$S = \{\text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 18}, \text{Math 109}, \text{Math 184}\}.$$

   There are two minimal elements: Math 20A and Math 18. Let us remove Math 18 from S and add it to the resulting list R.

2. Now we have
   $$R = \text{Math 18}$$
   and

   $$S = \{\text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 109}, \text{Math 184}\}.$$

   There is only one minimal element Math 20A. We remove it and add it to the list R.

3. On this step
   $$R = \text{Math 18}, \text{Math 20A}$$

and

$$S = \{\text{Math 20B}, \text{Math 20C}, \text{Math 109}, \text{Math 184}\}.$$

Again there is only one minimal element: Math 20B.

4.
$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}$$

and

$$S = \{\text{Math 20C}, \text{Math 109}, \text{Math 184}\}.$$

There is only one minimal element: Math 20C.

5.
$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}, \text{Math 20C}$$

and

$$S = \{\text{Math 109}, \text{Math 184}\}.$$

There is only one minimal element: Math 109.

6. Finally,

$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 109}$$

and

$$S = \{\text{Math 184}\}.$$

There is only one minimal element: Math 184A.

As a result, the final list is

$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 109}, \text{Math 184A}.$$

End of The Chapter Exercises

7.5 Show that the relation $|$ does not define a partial ordering on $\mathbb{Z}$.

7.6 Let a relation R be defined on the set of real numbers as follows: xRy iff $2x + y = 3$. Show that it is antisymmetric.

7.7 Are there any minimal elements in $\mathbb{N}$ with respect to $|$? Are there any maximal elements?

Part II

Introduction to Combinatorics

8 Bijections, Surjections, and Injections

Bijections, Surjections, and Injections: Introduction to Combinatorics #1

youtu.be/fW5Zxg0TMDc

In the previous chapters we used the property that the set is finite. However, we have never defined formally what it means. In this chapter we define cardinality which is a formalization of the notion size of the set and explain how to compare sizes of two sets.

8.1 Bijections

The simplest way to explain that one set has the same number of elements as another is to show a correspondence between elements of these sets. For example, in order to explain that the set $\{0, \pi, 1/4\}$ has the same number of elements as $\{1, 2, 3\}$ we may just say that 0 corresponds to 1, π corresponds to 2, and $1/4$ corresponds to 3. More formally such a correspondence is defined using the following definition.

Definition 8.1. *Let $f : X \to Y$ be a function. We say that f is a bijection iff the following properties are satisfied.*

- *Every element of Y is an image of some element of X. In other words,*

$$\forall y \in Y\ \exists x \in X\ f(x) = y.$$

- *Images of any two elements of X are different. In other words,*

$$\forall x_1, x_2 \in X\ f(x_1) \neq f(x_2).$$

Let us consider the following example. Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(x) = x + 1$; Note that it is a bijection:

- For any $y \in \mathbb{R}$, $f(y-1) = (y-1) + 1 = y$.
- If $f(x_1) = f(x_2)$, then $x_1 + 1 = x_2 + 1$ i.e. $x_1 = x_2$.

Exercise 8.1. *Show that x^3 is a bijection.*

One may notice that if we have a bijection f from $[n]$ to a set S we enumerate all the elements of S: $f(1), \ldots, f(n)$. This observation allows us to define the cardinality of a set.

Definition 8.2. *Let S be a set, we say that cardinality of S is equal to n (we write that $|S| = n$) iff there is a bijection from $[n]$ to S.*

We also say that a set T is finite if there is an integer n such that $|T| = n$.

Note that this definition does not guarantee that cardinality is unique so we need the following theorem.

Theorem 8.1. *For any set S, if there are bijections $f : [n] \to S$ and $g : [m] \to S$, then $n = m$.*

Before we prove this theorem, let us study some properties of bijections.

One of the nicest properties of bijections is that composition of two bijections is a bijection.

Theorem 8.2. *Let X, Y, and Z be some sets and $f : X \to Y$ and $g : Y \to Z$ be bijections. Then $(g \circ f) : X \to Z$ is also a bijection.*

Proof. We need to check two properties.

- Let $x_1 \neq x_2 \in X$. Note that $f(x_1) \neq f(x_2)$ since f is a bijection. Hence, $g(f(x_1)) \neq g(f(x_2))$ since g is a bijection as well. As a result, $(g \circ f)(x_1) \neq (g \circ f)(x_2)$.
- Let $z \in Z$; we need to find $x \in X$ such that $(g \circ f)(x) = z$. Note that since g is a bijection there is $y \in Y$ such that $g(y) = z$. Additionally, there is $x \in X$ such that $f(x) = y$ since f is a bijection. Thus, $(g \circ f)(x) = g(f(x)) = z$.

□

Probably the most important property of a bijection is that we may invert it.

Theorem 8.3. *Let $f : X \to Y$ be a function. f is invertible (i.e. there is a function $g : Y \to X$ such that $(f \circ g)(y) = y$ and $(g \circ f)(x) = x$ for all $x \in X$ and $y \in Y$) iff f is a bijection.*

Proof. $\Rightarrow$ Let's assume that f is invertible. We need to prove that f is a bijection.

- Let's assume that f does not satisfy the first property in the definitions of bijections i.e. there are $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$ but $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$, which is a contradiction.
- Let $y \in Y$. Note that $f(g(y)) = y$, hence, $\mathrm{Im} f = Y$.

$\Leftarrow$ Let's assume that f is bijective. We need to define a function $g : Y \to X$ which is an inverse of f. Let $y \in Y$, note that there is a unique x such that $f(x) = y$, we define $g(y) = x$. Note that

$f(g(y)) = y$ for every y by the construction of g. Additionally, $g(f(x)) = x$ since $f(g(f(x))) = f(x)$ and f is a bijection.

□

We denote g from this theorem as f^{-1} and in case when f is not a bijection $f^{-1}(y)$ denotes the set $\{x \in X \ : \ f(x) = y\}$.

Proof of Theorem 8.1. Let us consider the inverse g^{-1} of g (it exists by Theorem 8.3 since g is a bijection). Note that $h = g^{-1} \circ f$ is a bijection from $[n]$ to $[m]$.

We prove using induction by n that for any $n, m \in \mathbb{N}$, if there is a bijection h' from $[n]$ to $[m]$, then $n = m$. The base case is for $n = 1$; if $m \geq 2$, then there are $x, y \in [1]$ such that $h'(x) = 1$ and $h'(y) = 2$, but $x \neq y$ and we have only one element in $[1]$.

The induction step is also simple. Assume that there is a bijection h' from $[n+1]$ to $[m]$. We define a function $h'' : [n] \to [m-1]$ as follows:

$$h''(i) = \begin{cases} h'(i) & \text{if } h'(i) < h(n+1) \\ h'(i) - 1 & \text{otherwise} \end{cases}.$$

We prove that h'' is a bijection.

- Let $i_1 \neq i_2 \in [n]$. If $h'(i_1), h'(i_1) < h'(n+1)$ or $h'(i_1), h'(i_1) \geq h'(n+1)$, then $h''(i_1) \neq h''(i_2)$ since $h'(i_1) \neq h'(i_2)$. Otherwise, without loss of generality we may assume that $h'(i_1) < h(n+1) < h'(i_2)$ but it implies that $h''(i_1) = h'(i_1) < h'(n+1) \leq h'(i_2) - 1 = h''(i_2)$.

- Let $j \in [m-1]$. We need to consider two cases.

  1. Let $j < h(n+1)$. There is $i \in [n+1]$ such that $h'(i) = j$ since h' is a bijection (note that $i \neq n+1$). Thus $h''(i) = j$.
  2. Otherwise, there is $i \in [n+1]$ such that $h'(i) = j+1$ since h' is a bijection (note that $i \neq n+1$). Thus $h''(i) = j$.

Since h'' is a bijection, the induction hypothesis implies that $n = m - 1$. As a result, $n + 1 = m$. □

Using Theorem 8.3 we may notice that nonetheless that $X \times (Y \times Z)$ is not the same as $(X \times Y) \times Z$, there is a natural correspondence between the elements of these sets.

Theorem 8.4. *Let X, Y, Z be some sets. There are bijections from $X \times (Y \times Z)$ and $(X \times Y) \times Z$ to $\{(x, y, z) \ : \ x \in X, y \in Y, z \in Z\}$.*

Proof. Since the statement is symmetric, it is enough to prove that there is a bijection $f : X \times (Y \times Z) \to \{(x, y, z) \ : \ x \in X, y \in Y, z \in Z\}$. Define f such that $f(x, (y, z)) = (x, y, z)$. Clearly, $f^{-1}(x, y, z) = (x, (y, z))$ is the inverse of f, so f is indeed a bijection. □

Due to this correspondence we will think about elements $(x,(y,z))$, $((x,y),z)$, and (x,y,z) as they are equal to each other.

Also, using Theorem 8.3 we may finally prove that if there is a bijection from a finite set X to a finite set Y, then they have the same cardinality (i.e. they have the same number of elements).

Theorem 8.5. *Let X and Y be two finite sets such that there is a bijection f from X to Y. Then $|X| = |Y|$.*

Proof. Let $|X| = n$, and $g : [n] \to X$ be a bijection. Note that $f \circ g : [n] \to Y$ is a bijection, hence $|Y| = n$. □

Using this result we can make prove the following equality.

Corollary 8.1. *Let X be a finite set of cardinality n. Then 2^X has the same cardinality as $\{0,1\}^{|X|}$.*

Proof. To prove this statement we need to construct a bijection from 2^X to $\{0,1\}^{|X|}$. Let $|X| = n$ and $f : [n] \to X$ be a bijection.

First we construct a bijection $g_1 : 2^X \to 2^{[n]}$:

$$g_1(Y) = \{f(x) \;:\; x \in Y\}.$$

It is easy to see that the function

$$g_1^{-1}(Y) = \left\{f^{-1}(x) \;:\; x \in [n]\right\}$$

is an inverse of g_1, so g_1 is indeed a bijection.

Now we need to construct a bijection g_2 from $2^{[n]}$ to $\{0,1\}^n$: $g_2(Y) = (u_1, \ldots, u_n)$, where $u_i = 1$ iff $i \in Y$. It is clear that $g_2^{-1}(u_1, \ldots, u_n) = \{i \in [n] \;:\; u_i = 1\}$ is an inverse of g_2 so g_2 is indeed a bijection.

As a result, by Theorem 8.2, the function $(g_2 \circ g_1) : 2^X \to \{0,1\}^{|X|}$ is a bijection. □

8.2 Surjections and Injections

It is possible to note that the definition of the bijection consists of two part. Both of these parts are interesting in their own regard, so they have their own names.

Definition 8.3. *Let $f : X \to Y$ be a function.*

- *We say that f is a surjection iff every element of Y is an image of some element of X. In other words,*

$$\forall y \in Y \; \exists x \in X \; f(x) = y.$$

- *We say that f is an injection iff images of any two elements of X are different. In other words,*

$$\forall x_1, x_2 \in X \; f(x_1) \neq f(x_2).$$

Remark 8.1. *Let $f : X \to Y$ be an injection. Then $g : X \to \mathrm{Im} f$ such that $f(x) = g(x)$ is a bijection.*

Exercise 8.2. *Let $\mathbb{R}^+ = \{x \in \mathbb{R} \ : \ x > 0\}$. Is $f : \mathbb{R}^+ \to \mathbb{R}^+$ such that $f(x) = x + 1$ a surjection/injection?*

Like in the case of the bijection we may use surjections and injections to compare sizes of sets.

Theorem 8.6. *Let X and Y be finite sets.*

- *If there is an injection from X to Y, then $|X| \leq |Y|$.*
- *If there is a surjection from X to Y, then $|X| \geq |Y|$.*

8.3 Generalized Commutative Operations

Using the notation of cardinality we may generalize the summation operation in the following way: $\sum_{i \in S \ : \ P(i)} f(i)$ is equal to the sum of $f(i)$ for all the $i \in \{i \in S \ : \ P(i)\}$; i.e.

$$\sum_{i \in S \ : \ P(i)} f(i) = \sum_{j=1}^{k} f(i_j),$$

where $\{i \in S \ : \ P(i)\} = \{i_1, \ldots, i_k\}$. More formally,

$$\sum_{i \in S \ : \ P(i)} f(i) = \sum_{j=1}^{k} f(g(j)),$$

where $k = |\{i \in S \ : \ P(i)\}|$ and $g : \{i \in S \ : \ P(i)\} \to [k]$ is a bijection.

Theorem 8.7. *The definition of $\sum_{i \in S : P(i)} f(i)$ is correct; i.e. $\sum_{i=1}^{k} f(g_1(i)) = \sum_{i=1}^{k} f(g_2(i))$ for any two bijections $g_1, g_2 : \{i \in S \ : \ P(i)\} \to [k]$.*

Proof. Proof of this theorem consists of two parts. First, we prove that

$$\sum_{i=1}^{k} f(g(i)) = \sum_{i=1}^{k} f(g(h(i))) \tag{8.1}$$

for any bijections $g : \{i \in S \ : \ P(i)\} \to [k]$ and $h : [k] \to [k]$.

To prove this statement, we introduce the notion of inversion. We say that $i, j \in [k]$ for an inversion in h iff $h(i) > h(j)$ and $i < j$. We denote by $I(h)$ the number of inversions in h; i.e. $I(h) = |\{(i, j) \ : \ i, j \text{ form an inversion in } h\}|$. It is easy to see that $I(h) = 0$ iff $h(i) = i$ for any $i \in [k]$. It is also clear that if i, j form an inversion in h, then $I(h) > I(h')$, where

$$h'(x) = \begin{cases} h(j) & \text{if } x = i \\ h(i) & \text{if } x = j \\ h(x) & \text{otherwise} \end{cases}.$$

We prove Equation 8.1 using the induction by $I(h)$.

(the base case) If $I(h) = 0$, then h is an identity function and $g(i) = g(h(i))$. Hence, Equation 8.1 is true.

(the induction step) By the induction hypothesis, if $I(h') < \ell$, then

$$\sum_{i=1}^{k} f(g(i)) = \sum_{i=1}^{k} f(g(h'(i)))$$

for any bijection $h' : [k] \to [k]$. Let us consider a bijection $h : [k] \to [k]$ such that $I(h) = \ell$. Define $h' : [k] \to [k]$ such that

$$h'(x) = \begin{cases} h(j) & \text{if } x = i \\ h(i) & \text{if } x = j \\ h(x) & \text{otherwise} \end{cases}.$$

Note that by the induction hypothesis,

$$\sum_{i=1}^{k} f(g(i)) = \sum_{i=1}^{k} f(g(h'(i)))$$

and it is clear that

$$\sum_{i=1}^{k} f(g(h'(i))) = \sum_{i=1}^{k} f(g(h(i))).$$

As a result, Equation 8.1 is true.

Now we are ready to finish proof of the theorem. Consider $g_1, g_2 : \{i \in S \ : \ P(i)\} \to [k]$ and define $h = g_1^{-1} \circ g_2$. Note that $h : [k] \to [k]$ is a bijection and $g_1(h(i)) = g_2(i)$. Thus we proved that

$$\sum_{i=1}^{k} f(g_1(i)) = \sum_{i=1}^{k} f(g(h(i))) = \sum_{i=1}^{k} f(g_2(i)).$$

□

Similarly one may define a generalized union and intersection of sets. Let Ω and S be some sets, $X : S \to 2^{\Omega}$ and $P(i)$ be a predicate. Then

$$\bigcup_{i \in S \ : \ P(i)} X(i) = \bigcup_{i=1}^{k} X(g(i))$$

and

$$\bigcap_{i \in S \ : \ P(i)} X(i) = \bigcap_{i=1}^{k} X(g(i)),$$

where $k = |\{i \in S \ : \ P(i)\}|$ and $g : \{i \in S \ : \ P(i)\} \to [k]$ is a bijection.

Exercise 8.3. *Show that the definitions of $\bigcup_{i \in S : P(i)} X(i)$ and $\bigcap_{i \in S : P(i)} X(i)$ are correct, i.e. that they do not depend on the choice of g.*

End of The Chapter Exercises

8.4 Construct a bijection from $\{0, 1, 2\}^n$ to

$$\{(A, B) \ : \ A, B \subseteq [n] \text{ and } A, B \text{ are disjoint}\} .$$

8.5 Construct a bijection from $\{0, 1\} \times [n]$ to $[2n]$.

8.6 Prove Theorem 8.6.

9 Counting Principles

Counting Principles:
Introduction to Combinatorics #2

youtu.be/dAoperLCjb8

9.1 The Additive Principle

The first principle is called *additive* principle and it states that if you have two disjoint sets, then their union have size equal to the sum of their sizes.

A simple illustration of this statement is the following. Assume you have three pencils and two pens; how many ways to choose a writing accessory. According to this principle the answer is $2 + 3 = 5$.

Theorem 9.1 (The Additive Principle)**.** *Let X and Y be finite sets. If $X \cap Y = \emptyset$, then $|X \cup Y| = |X| + |Y|$.*

Proof. Let $|X| = n$, $|Y| = m$ and $g : [n] \to X$ and $h : [m] \to Y$ be bijections. In order to prove it we just construct a bijection $f : [n+m] \to (X \cup Y)$.

$$f(i) = \begin{cases} g(i) & i < n \\ h(i-n) & i > n \end{cases}.$$

It's easy to see that f is an injection. Let us start by assuming the opposite i.e. that $i_0 \neq i_1 \in X \cup Y$ such that $f(i_0) = f(i_1)$. There are three cases.

- The first is when $i_0, i_1 \in [n]$. In this case $g(i_0) = g(i_1)$ which contradicts the assumption that g is a bijection.
- The second is when $i_0, i_1 \in \{n+1, n+2, \ldots, m\}$. In this case $h(i_0 - n) = h(i_1 - n)$ which contradicts the assumption that h is a bijection.
- Finally, the last case is when $i_0 \in [n]$ and $i_1 \in \{n+1, n+2, \ldots, m\}$. It is easy to see that this implies that $g(i_0) = h(i_1 - n)$. However, it means that $g(i_0) = h(i_1 - n) \in (X \cap Y)$, which contradicts the assumption that $X \cap Y = \emptyset$.

To finish the proof we need to show that f is a surjection. Let $w \in (X \cup Y)$. Consider the following two cases.

- Let $w \in X$. There is $i \in [n]$ such that $f(i) = g(i) = w$ since g is a bijection.

- Otherwise, $w \in Y$. In this case, there is $i \in [m]$ such that $f(i+n) = h(i) = w$ since h is a bijection.

□

Corollary 9.1. *Let $X_1, \ldots, X_n$ be some pairwise disjoint sets. Then $|\bigcup_{i=1}^n X_i| = \sum_{i=1}^n |X_i|$.*

Exercise 9.1. *Prove Corollary 9.1.*

9.2 *The Multiplicative Principle*

The next principle is called the *multiplicative* principle and it can be illustrated as follows: imagine that you are given two postal stamps and three envelopes, how many ways are there to pack the letters? The answer is obviously $2 \cdot 3 = 6$.

Theorem 9.2 (The Multiplicative Principle)**.** *Let X and Y be finite sets. Then $|X \times Y| = |X| \times |Y|$.*

Proof. If one of the sets X and Y is empty, then $X \times Y$ is empty as well and the statement is as follows.

Assume that none of the sets are empty. Let $|X| = n$, $|Y| = m$, and $f : [n] \to X$ and $g : [m] \to Y$ be bijections. Note that

$$\bigcup_{i=1}^{n} (\{f(i)\} \times Y) = X \times Y.$$

Additionally, note that $(\{f(i)\} \times Y) \cap (\{f(j)\} \times Y) = \varnothing$ for $i \neq j$. Finally, it is easy to see that $g_i : [m] \to (\{f(i)\} \times Y)$ such that $g_i(j) = (f(i), g(j))$ is a bijection. Hence, $|X \times Y| = \sum_{i=1}^n |\{f(i)\} \times Y| = n \cdot m$. □

Exercise 9.2. *Find the cardinality of the set*

$$\{(x, y) \ : \ x, y \in [9] \text{ and } x \neq y\}.$$

By analogy with unions and intersections of many sets we can define the cross product of many sets. Let $X_1, \ldots, X_n$ be some sets. Then $\bigtimes_{i=1}^1 X_i = A_1$ and $\bigtimes_{i=1}^{k+1} X_i = \left(\bigtimes_{i=1}^k X_i\right) \times X_{k+1}$[1].

[1] Note that cross product is not associative and different definitions of the product of several sets are not equivalent. However, the bijection constructed in the previous section allow us to think about these definitions as if they are equivalent.

Corollary 9.2. *Let $X_1, \ldots, X_n$ be some finite sets. Then $|\bigtimes_{i=1}^n X_i| = \prod_{i=1}^n |X_i|$.*

Exercise 9.3. *Prove Corollary 9.2.*

Theorem 9.3. *For any set X, $|2^X| = 2^{|X|}$.*

Proof. By Corollary 8.1, $|2^X| = \left|\{0,1\}^{|X|}\right|$, so it is enough to prove that $|\{0,1\}^{|X|}| = 2^{|X|}$. This statement is true by Corollary 9.2 since $|\{0,1\}^{|X|}| = \prod_{i=1}^{|X|} |\{0,1\}| = 2^{|X|}$. □

9.3 The Inclusion-exclusion Principle

The last principle we are going to discuss in this chapter is the inclusion-exclusion principle which helps us to find the size of the union of sets when they are not disjoint.

Theorem 9.4 (The Inclusion-exclusion Principle)**.** *Let X and Y be finite sets. Then $|X \cup Y| = |X| + |Y| - |X \cap Y|$.*

Proof. Note that $X \cup Y = (X \setminus Y) \cup (Y \setminus X) \cup (X \cap Y)$. Hence, $|X \cup Y| = |X \setminus Y| + |Y \setminus X| + |X \cap Y|$. But it is possible to note that $|Y \setminus X| + |X \cap Y| = |Y|$ and $|X \setminus Y| + |X \cap Y| = |X|$. □

Corollary 9.3. *Let $X_1, \ldots, X_n$ be some finite sets. Then*

$$\left|\bigcup_{i=1}^{n} X_i\right| = \sum_{S \subseteq [n]\,:\, S \neq \varnothing} (-1)^{|S|+1} \left|\bigcap_{i \in S} X_i\right|.$$

Proof. As always, we prove this statement using induction by n. The base case for $n = 2$ is true by Theorem 9.4.

By the induction hypothesis,

$$\left|\bigcup_{i=1}^{k} X_i\right| = \sum_{S \subseteq [k]\,:\, S \neq \varnothing} (-1)^{|S|+1} \left|\bigcap_{i \in S} X_i\right|.$$

In addition, by Theorem 9.4,

$$\left|\bigcup_{i=1}^{k+1} X_i\right| = \left|\bigcup_{i=1}^{k} X_i\right| + |X_{k+1}| - \left|\left(\bigcup_{i=1}^{k} X_i\right) \cap X_{k+1}\right|.$$

We need to simplify two elements of the sum on the right of the equality. By the induction hypothesis,

$$\left|\bigcup_{i=1}^{k} X_i\right| = \sum_{S \subseteq [k]\,:\, S \neq \varnothing} (-1)^{|S|+1} \left|\bigcap_{i \in S} X_i\right|.$$

In addition, it is easy to note that

$$\left|\left(\bigcup_{i=1}^{k} X_i\right) \cap X_{k+1}\right| = \left|\bigcup_{i=1}^{k} (X_i \cap X_{k+1})\right|.$$

Thus using the induction hypothesis,

$$\left|\left(\bigcup_{i=1}^{k} X_i\right) \cap X_{k+1}\right| =$$

$$\sum_{S \subseteq [k]\,:\, S \neq \varnothing} (-1)^{|S|+1} \left|\bigcap_{i \in S} (X_i \cap X_{k+1})\right| =$$

$$\sum_{S \subseteq [k+1]\,:\, (k+1) \in S \text{ and } S \neq \{k+1\}} (-1)^{|S|} \left|\bigcap_{i \in S} X_i\right|.$$

As a result,

$$|X_{k+1}| - \left|\left(\bigcup_{i=1}^{k} X_i\right) \cap X_{k+1}\right| = \sum_{S\subseteq[k+1]\,:\,(k+1)\in S} (-1)^{|S|+1} \left|\bigcap_{i\in S} X_i\right|.$$

Which implies that

$$\left|\bigcup_{i=1}^{k+1} X_i\right| = \sum_{S\subseteq[k]\,:\,S\neq\emptyset} (-1)^{|S|+1} \left|\bigcap_{i\in S} X_i\right| + \sum_{S\subseteq[k+1]\,:\,(k+1)\in S} (-1)^{|S|+1} \left|\bigcap_{i\in S} X_i\right| = \sum_{S\subseteq[k+1]\,:\,S\neq\emptyset} (-1)^{|S|+1} \left|\bigcap_{i\in S} X_i\right|.$$

□

End of The Chapter Exercises

9.4 How many numbers from [999] are not divisible neither by 3, nor by 5, nor by 7.

9.5 How many numbers x from 1 to 999 such that at least one of the digits of x is 7?

9.6 Let A, B be some finite sets such that $A \subseteq B$. Show that $|B \setminus A| = |B| - |A|$.

9.7 Let n be some positive integer. Find the cardinality of the set

$$\{(A, B)\ :\ A, B \subseteq [n] \text{ and } A \cap B \neq \emptyset\}?$$

10 The Pigeonhole Principle

The Pigeonhole Principle: Introduction to Combinatorics #3

youtu.be/1D1Fa7WIUO8

The principle we are going to discuss in this chapter is very simple, it states that if you have more objects than boxes, then you cannot put all the objects into boxes without puting two objects into the same box.

More formally the principle can be formulated as follows: if $n > m$, then any function from $[n]$ to $[m]$ is not an injection. This simple statement is famous in mathematics and called *the pigeonhole principle*[1].

[1] The pigeonhole principle is also called the Dirichlet principle, after the German mathematician G. Lejeune Dirichlet, who demonstrated, using this principle, that there were at least two Parisians with the same number of hairs on their heads.

Theorem 10.1 (the pigeonhole principle). *Let X and Y be some sets such that $|X| > |Y|$. Then for any function $f : X \to Y$ there are $x_0 \neq x_1 \in X$ such that $f(x_0) = f(x_1)$.*

Proof. The statement follows from Theorem 8.6. □

This simple statement is very handy in combinatorics. For example, using this statement one may prove that in any group of more than 12 people there are two people who were born in the same month.

Assume that there are n people in the group and $n > 12$. Consider the following function $f : [n] \to [12]$ such that $f(i) = j$ if the ith person was born in jth month. Note that f is not an injection since $n > 12$ i.e. there are $i_0 \neq i_1$ such that i_0th and i_1th person are born in the same month.

We may also prove that in any group of people there are two people who are friends with the same number of people in the group.

Assume the number of people is n. It is easy to see that every person may have at most $n - 1$ friends. Hence, we may define a function $f : [n] \to \{0, \dots, n-1\}$ such that $f(i)$ is equal to the number of friends in this group of the ith person in this group. We need to consider two cases.

- If $\mathrm{Im} f \subseteq [n-1]$, then $|[n]| > |\mathrm{Im} f|$ and f is not an injection.
- Otherwise, note that it is not possible that $(n-1) \in \mathrm{Im} f$ because if there is a friend with no friends it is not possible that there is a friend who is friends with everyone. Hence, $f : [n] \to \{0, 1, \dots, n-2\}$ and f is not an injection.

Theorem 10.2 (Erdős—Szekeres). *Every sequence of $(r-1)(s-1)+1$ distinct real numbers contains a subsequence of length r that is increasing or* *a subsequence of length s that is decreasing.*

Proof. Given a sequence of length $(r-1)(s-1)+1$, label each number x_i in the sequence with the pair (a_i, b_i), where a_i is the length of the longest increasing subsequence ending with x_i and b_i is the length of the longest decreasing subsequence ending with x_i. Each two numbers in the sequence are labeled with a different pair: if $i < j$ and $x_i < x_j$ then $a_i < a_j$, and on the other hand if $x_i > x_j$ then $b_i < b_j$. But there are only $(r-1)(s-1)$ possible labels if a_i is at most $r-1$ and b_i is at most $s-1$, so by the pigeonhole principle there must exist a value of i for which a_i or bi is outside this range. If ai is out of range then x_i is part of an increasing sequence of length at least r, and if b_i is out of range then x_i is part of a decreasing sequence of length at least s. □

10.1 The Generalized Pigeonhole Principle

One may generalize the pigeonhole principle in the following way. If N objects are placed into k boxes, then there is at least one box containing at least $\lceil N/k \rceil$ objects.

Theorem 10.3 (the generalized pigeonhole principle)**.** *Let X and Y be some sets. Then for any function $f : |X| \to |Y|$ there are $x_1, \ldots, x_\ell \in X$ such that*

- $f(x_i) = f(x_j)$,
- $x_i \neq x_j$ *for any* $i \neq j \in [\ell]$, *and*
- $\ell \geq \lceil |X|/|Y| \rceil$, *where* $\lceil \alpha \rceil$ *denotes the least integer greater than or equal to* α.

Now we illustrate applications of this principle on some examples and prove the statement in the next section.

Using this theorem we can prove that if we draw 9 cards out of a deck of cards, we are guaranteed that at least three of them are of the same suit. Given that, there are 4 suits in the deck, by pigeonhole principle if we put each card into one of the four boxes according to their suits, one of the boxes should have at least $\lceil 9/4 \rceil = 3$ cards.

Another example shows how the generalized pigeonhole principle can be applied to an important part of combinatorics called Ramsey theory.

Assume that in a group of six people, each pair of individuals consists of two friends or two enemies. One may prove that there are either three mutual friends or three mutual enemies in the group.

Let A be one of the six people; of the five other people in the group, there are either three or more who are friends of A, or three or more who are his enemies A. This statements follows from the generalized pigeonhole principle since when five objects are divided into two sets, one of the sets has at least $\lceil 5/2 \rceil = 3$ elements. Without loss of generality we may suppose that B, C, and D are friends of A. If any two of these three individuals are friends, then these two and A form a group of three mutual friends. Otherwise, B, C, and D form a set of three mutual enemies.

10.2 The Averaging Principle

Assume that we have a collection of m objects, the ith of which has "size" l_i. We wish to show that at least one of the objects is large. In this situation we can argue that at least one of the objects has size greater or equal to the average size ($\sum l_i/m$).

Theorem 10.4 (the averaging principle)**.** *Every sequence of numbers has a number at least as large as the average and a number at least as small as the average; i.e. for any sequence $a_1, \ldots, a_m$ there are i and j such that*

$$a_i \geq \frac{1}{m}\sum_{i=1}^{m} a_i$$

and

$$a_j \leq \frac{1}{m}\sum_{i=1}^{m} a_i.$$

Proof. We prove only the existence of i, proof of the existence of j is almost the same.

Assume the opposite, i.e. that $a_i < \sum_{i=1}^{n} a_i/m$ for any $i \in [n]$. Note that this implies that $\sum_{i=1}^{n} a_i \leq m \cdot \sum_{i=1}^{n} a_i/m = \sum_{i=1}^{n} a_i$. Which is a contradiction. $\square$

Exercise 10.1. *Finish the proof of Theorem 10.4*

Like the pigeonhole principle, this principle is very simple but the applications of it are surprisingly intresting.

First, it allows to prove the generalized pigeonhole principle.

Proof of Theorem 10.3. Let $Y = [m]$ (it is easy to see that the proof works for any other finite Y). Define the sequence $a_i = |f^{-1}(i)|$. Note that we need to prove that $a_i \geq \lceil |X|/m \rceil$ for some $i \in [m]$

It is clear that $\bigcup_{i=1}^{m} f^{-1}(i) = X$ and that $f^{-1}(i) \cap f^{-1}(j) = \emptyset$ for any $i \neq j \in [m]$. Thus, by the additive principle, $\sum_{i=1}^{m} a_i = |X|$. Hence, by the averaging principle, $a_i \geq |X|/m$ for some $i \in [m]$. However, a_i is an integer, thus $a_i \geq \lceil |X|/m \rceil$. $\square$

Another nice application of the averaging principle allows us to prove that if in some group (with more than one person) the number of pairs of people who know each other is less than $n-1$, then we can split this group into two subgroups such that people from different subgroups do not know each other.

Let us assume that there are n people in the group. We prove the statement using the induction by n.

(the base case) If $n = 2$, there are less than $n-1 = 1$ pairs of people who know each other, in other words, these two people in the group do not know each other. Thus we can put each of them into a separate subgroup.

(the induction step) Let p_i ($i \in [n]$) be the number of acquaintances of the ith person. Note that $\sum_{i=1}^{n} p_i \leq 2(n-2)$ since we count each pair twice. By the averaging principle, $p_i \leq 2(n-2)/n = 2 - 2/n$ for some $i \in [n]$. Thus p_i is either 0 or 1.

- If $p_i = 0$, we can put the ith person into the first subgroup and everyone else into another.
- If $p_i = 1$ we consider the group of $n-1$ people without the ith person, by the induction hypothesis, we can split everyone but ith person into two subgroups and since the ith person has only one acquaintance we can put them in the same subgroup.

End of The Chapter Exercises

10.2 Show that among any group of five (not necessarily consecutive) integers, there are two with the same remainder when divided by 4.

10.3 Show that if there are 30 students in a class, then at least two have last names that begin with the same letter.

10.4 Let n be a positive integer. Show that in any set of n consecutive integers there is exactly one divisible by n.

10.5 Prove that for every integers $a_1, \ldots, a_n$ there are $k > 0$ and $\ell \geq 0$ such that $k + \ell \leq n$ and $\sum_{i=k}^{k+\ell} a_i$ is divisible by n.

10.6 Let $S \subseteq [20]$ be a set. Show that if $|S| \geq 13$, then there are $a, b \in S$ such that $a - b = 6$.

10.7 How many numbers must be selected from the set [6] to guarantee that at least one pair of these numbers add up to 7?

10.8 Sasha is training for a triathlon. Over a 30 day period, he pledges to train at least once per day, and 45 times in all. Then there will be a period of consecutive days where he trains exactly 14 times.

10.9 Show that among any $n+1$ positive integers not exceeding $2n$ there must be an integer that divides one of the other integers. *Hint: Consider the set of holes equal to the set of odd numbers from* 1 *to* $2n$.

10.10 Let a_1, a_2, …, a_t be positive integers. Show that if $a_1 + a_2 + \cdots + a_t - t + 1$ objects are placed into t boxes, then for some $i \in [t]$, the ith box contains at least a_i objects. *Hint: It is important in this question that* $a_1, \ldots, a_t$ *are integers.*

10.11 Let $\{(x_1, y_1), \ldots, (x_5, y_5)\} \subseteq \mathbb{Z}^2$ be a set of five distinct points with integer coordinates in the xy plane. Show that the midpoint of the line joining at least one pair of these points has integer coordinates.

11 Binomial Coefficients

Permutations and Binomial Coefficients: Introduction to Combinatorics #4

`youtu.be/HLClazoqqzg`

11.1 Counting Functions

Assume we have two finite sets X and Y. The first question we may ask is how many function exist from X to Y?

Theorem 11.1. *Let X and Y be some finite sets. Y^X represents the set of all functions from X to Y. Then $|Y^X| = |Y|^{|X|}$.*

Proof. For simplicity we prove the statement in the case when $X = [n]$. Fix some finite set Y. We prove the statement using induction by n. The base case for $n = 1$ is obvious, since there are $|Y|$ different functions from $[1]$ to Y. Let us prove the induction step, by the induction hypothesis, $|Y^{[n-1]}| = |Y|^{n-1}$. Note that

$$|Y^{[n]}| = \left|\left\{(f, y) \ : \ f \in Y^{[n-1]}, y \in Y\right\}\right| = \\ |Y^{[n-1]} \times Y| = |Y^{[n-1]}| \cdot |Y| = |Y|^n.$$

□

Exercise 11.1. *Finish the proof of Theorem 11.1 by proving that the statement holds for any set X.*

However, what if we need to find size of a subset of Y^X satisfying some constraint? For example, we may try to find the size of the set

$$(Y)_X = \left\{f \in Y^X \ : \ f \text{ is an injection}\right\}.$$

First, let us try to do this informally. Assume that $X = [n]$ and $|Y| = m$, to define $f \in (Y)_X$ we need to choose images of $1, 2, \ldots, n$. There are m possible ways to select an image of 1, $m - 1$ ways to define $f(2)$ since we cannot use the value selected for 1 etc. Hence, $|(Y)_X| = m(m-1)\ldots(m-n+1)$ (we denote this number as $(m)_n$).

Theorem 11.2. *Let X and Y be some sets. Then $|(Y)_X| = (|Y|)_{|X|}$.*

Proof. Let us prove this statement for $X = [n]$. We prove this using induction by n. The base case, for $n = 1$, is clear. Now we need to prove the induction step from n to $n+1$. By the induction hypothesis, for any m, the number of injections from $[n]$ to Y is equal to $(|Y|)_n$.

Fix some m and some set Y of cardinality m. Note that

$$|(Y)_X| = |\left\{(f,v) \in (Y)_{[n-1]} \times [m] \ : \ v \notin \mathrm{Im} f\right\}|.$$

It is easy to see that $|\{(f,v) \ : \ v \notin \mathrm{Im} f\}| = m-n+1$ for any $f \in (Y)_{[n-1]}$ and

$$\left\{(f,v) \in (Y)_{[n-1]} \times [m] \ : \ v \notin \mathrm{Im} f\right\} = \bigcup_{f \in (Y)_{[n-1]}} \{(f,v) \ : \ v \notin \mathrm{Im} f\}.$$

As a result, $|(Y)_X| = (m)_{n-1} \cdot (m-n+1) = (m)_n$. $\square$

The special case of this result is that there are $n \cdot (n-1) \cdot \ldots \cdot 1$ different bijections from $[n]$ to $[n]$. Such bijections are called permutations and the number is denoted by $n!$.

Exercise 11.2. *Finish the proof of Theorem 11.2 by proving that the statement holds for any set X.*

Let us consider the following problem: given a set of m different objects, how many ways to select n of them in some order (i.e. if the objects are numbers 1, 2, and 3, then selecting 1 and 2 is not the same as selecting 2 and 1). It is clear the the number of ways to select these n objects is the same as the number of injections f from $[n]$ to the set of objects since the first object is $f(1)$, the second is $f(2)$ etc. As a result,

Corollary 11.1. *There are $(m)_n$ ways to select n objects if the order matters.*

11.2 Counting Subsets

In this section we consider a modification of the previous problem. We count the number of ways to select n objects out of m if the order does not matter. Note that it is the same as counting the number of subsets of cardinality n of the set of m objects.

Recall that we denoted the set of all subsets of X by 2^X. The reason for this notation is that $|2^X| = 2^{|X|}$. A quite famous example of a subset of this set is the set

$$\binom{X}{k} = \{A \subseteq X \ : \ |A| = n\}.$$

In other words, it is the set of all possible ways to select n elements from X. Size of the set $\binom{[m]}{n}$ we denote by $\binom{m}{n}$ and call it a binomial coefficient.

Exercise 11.3. *Show that for any two finite sets X and Y, if $|X| = |Y|$, then $\left|\binom{X}{k}\right| = \left|\binom{Y}{k}\right|$.*

Note that by any ordered selection of n object out of m, one may construct an unordered selection of n objects out of m, and each unordered selection is counted $n!$.

Theorem 11.3. *For any $n > k \geq 0$, $\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}$.*

Exercise 11.4. *Show that $\binom{n}{k} = \binom{n}{n-k}$ for any $n > k$.*

The formula in the Thorem 11.3 allows to find the values of binomial ceofficients, however, it is not very convinient since $n!$ is growing very fast. Thus the following theorem provides a much more efficient way to compute the values of binomial coefficients.

Theorem 11.4 (Pascal's rule)**.** *For $n > k \geq 1$, $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.*

Proof. The first, algebraic, proof of this theorem is quite simple, we just notice that

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-k-1)!} = \\ \frac{(n-1)!}{(k-1)!(n-k-1)!}\left(\frac{1}{n-k} + \frac{1}{k}\right) = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

However, this proof does not explain *why* the statement is true. So we consider an alternative proof, which informally can be explained as follows. Assume we need to choose k objects out of n. There are two possible ways:

- we may select n and choose $k-1$ objects from the rest,
- or we may decide to not select n choose k objects from the rest.

In the first case we have $\binom{n-1}{k-1}$ ways to select objects and in the second case we have $\binom{n-1}{k}$ ways to select objects.

Let us prove the statement a bit more formally. Note that

$$\binom{[n]}{k} = \{A \subseteq [n] \ : \ |A| = k \text{ and } n \in A\} \cup \\ \{A \subseteq [n] \ : \ |A| = k \text{ and } n \notin A\}.$$

Since these sets are disjoint and $\{A \subseteq [n] \ : \ |A| = k \text{ and } n \notin A\} = \binom{[n-1]}{k}$, we get the following equality

$$\binom{n}{k} = |\{A \subseteq [n] \ : \ |A| = k \text{ and } n \in A\}| + \binom{n-1}{k}.$$

Hence, to finish the proof we need to explain that

$$|\{A \subseteq [n] \ : \ |A| = k \text{ and } n \in A\}| = \binom{n-1}{k-1}.$$

To prove this statement we construct a bijection

$$f : \{A \subseteq [n] \;:\; |A| = k \text{ and } n \in A\} \to \binom{[n-1]}{k}$$

such that $f(A) = A \setminus \{n\}$. It is clear that this is a bijection. Thus, we prove the statement. □

A mnemonic rule for the Pascal's rule is to use Pascal's triangle. [1]

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

In this diagram the $k$the entry of the nth row (entries and rows have numbers starting from 0) is equal to $\binom{n}{k}$. Thus the rule for the triangle is very simple, the value of an entry is equal to 1 if it is the first or the last in the row or it is equal to the sum of the two entries to the left and right on the row above.

Exercise 11.5. *Show that* $\binom{n}{k} = \binom{n}{n-k}$ *for any intgers* $n > k \geq 0$

Now we are ready to prove the theorem which gave the name to binomial coefficients.

Theorem 11.5 (Binomial theorem)**.** *For any real numbers* x *and* y,

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n.$$

Proof. Informally, the explanation of the equality is as follows. If we consider the product

$$\underbrace{(x+y) \cdot (x+y) \cdot \ldots \cdot (x+y)}_{n \text{ times}},$$

then for every k there are exactly $\binom{n}{k}$ possibilities to obtain $x^k y^{n-k}$. Indeed, to obtain $x^k y^{n-k}$ we need to choose x from n possibilities (corresponding to the multiplier $x+y$) exactly k times.

A formal proof uses the induction by n. The base case is true, since $\sum_{k=0}^{1} \binom{1}{k} x^k y^{1-k} = x + y = (x+y)^1$. Assume that

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n,$$

[1] The pattern of numbers that forms Pascal's triangle was known well before Pascal's time. Halayudha, around 975 explained obscure references to Meru-prastaara, the Staircase of Mount Meru, giving the first surviving description of the arrangement of these numbers into a triangle.

The Persian mathematician Al-Karaji (953–1029) wrote a now lost book which contained the first description of Pascal's triangle. It was later repeated by the Persian poet-astronomer-mathematician Omar Khayyám (1048–1131); thus the triangle is also referred to as the Khayyam triangle in Iran.

Pascal's triangle was known in China in the early 11th century through the work of the Chinese mathematician Jia Xian (1010–1070). In the 13th century, Yang Hui (1238–1298) presented the triangle and hence it is still called Yang Hui's triangle in China.

Pascal's Traité du triangle arithmétique (Treatise on Arithmetical Triangle) was published in 1655. In this, Pascal collected several results then known about the triangle, and employed them to solve problems in probability theory. The triangle was later named after Pascal by Pierre Raymond de Montmort (1708) who called it "Table de M. Pascal pour les combinaisons" (French: Table of Mr. Pascal for combinations) and Abraham de Moivre (1730) who called it "Triangulum Arithmeticum PASCALIANUM" (Latin: Pascal's Arithmetic Triangle), which became the modern Western name.

we wish to prove that

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} = (x+y)^{n+1}.$$

Note that

$$\begin{aligned}(x+y)^{n+1} = (x+y)\left(\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}\right) = \\ \sum_{k=0}^{n}\binom{n}{k}x^{k+1}y^{n-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n+1-k} = \\ \sum_{k=1}^{n+1}\binom{n}{k-1}x^k y^{n+1-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n+1-k} = \\ \sum_{k=0}^{n+1}\left(\binom{n}{k-1}+\binom{n}{k}\right)x^k y^{n+1-k} = \sum_{k=0}^{n+1}\binom{n+1}{k}x^k y^{n+1-k}.\end{aligned}$$

□

Counting Groups of Subsets

In this section we study a generalization of the question we study in the previous section: "How many ways to select ℓ groups made of k_1, k_2, …, k_ℓ objects, respectively, out of n". We denote this number by $\binom{n}{k_1\ k_2\ \ldots\ k_\ell\ (n-m)}$, where $m = k_1 + \cdots + k_\ell$.

Clearly selecting these objects is the same as selecting k_1 objects out of n, after that selecting k_2 objects out of $n - k_1$ etc. As a result,

$$\begin{aligned}\binom{n}{k_1\ k_2\ \ldots\ k_\ell\ (n-m)} = \\ \frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdot\ldots\cdot\frac{(n-k_1-k_2-\cdots-k_{\ell-1})!}{k_\ell!(n-k_1-k_2-\cdots-k_\ell)!} = \\ \frac{n!}{k_1!k_2!\ldots k_\ell!(n-k_1-k_2-\cdots-k_\ell)!}.\end{aligned}$$

Similarly to the Binomial theorem, we can prove the following.

Theorem 11.6 (Multinomial theorem)**.** *For any real numbers x_1, x_2, …, x_ℓ and intger n,*

$$(x_1 + x_2 + \cdots + x_\ell)^n = \sum_{k_1,k_2,\ldots,k_\ell\ :\ k_1+k_2+\cdots+k_\ell=n} \binom{n}{k_1\ k_2\ \ldots\ k_\ell} \prod_{i=1}^{n} x_i^{k_i}.$$

Exercise 11.6. *Prove Theorem 11.6.*

11.3 Double Counting

The method that was used to prove Theorem 11.4 can be generalized to a method that is caleed *double counting principle*. The double counting

Double Counting:
Introduction to Combinatorics #4

youtu.be/OrzMP8nuuho

principle states the following "obvious" fact: if the size of a set is counted in two different ways, the answers are the same.

Using this principle we may cprove the following theorem.

Theorem 11.7 (Vandermonde's identity)**.** *For any integers $n, m > k$, $\sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$.*

Proof. The idea is as follows, let us imagine that we have n parrots and m crows, and we need to find how many ways to select k birds. It is easy to see that it is equal to $\binom{n+m}{k}$. At the same if we need to select i parrots there are $\binom{n}{i}\binom{m}{k-i}$ ways to do this. Thus the number is also equal to $\sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}$. □

However, the method can be used in a more sophisticated way.

Lemma 11.1 (Handshaking Lemma)**.** *Suppose some number of people meet at a party and some shake hands. Assume that no person shakes his or her own hand and furthermore no two people shake hands more than once.*

The number of guests who shake hands an odd number of times is even.

Proof. Let $1, \ldots, n$ be the people at the party. We apply double counting to the set of ordered pairs (i, j) for which i and j shake hands with each other at the party. Let d_i be the number of times that i shakes hands, and e be the total number of handshakes that occur. On one hand, the number of pairs is $\sum_{i=1}^{n} d_i$, since for each i the number of choices of j is equal to d_i. On the other hand, each handshake gives rise to two pairs (i, j) and (j, i); so the total is $2e$. Thus $\sum_{i=1}^{n} d_i = 2e$. But, if the sum of n numbers is even, then evenly many of the numbers are odd. (Because if we add an odd number of odd numbers and any number of even numbers, the sum will be always odd). □

End of The Chapter Exercises

11.7 Show that $(x + y)_n = \sum_{k=0}^{n} \binom{n}{k} (x)_k (y)_{n-k}$.

11.8 Show that $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$.

11.9 Show that $\sum_{m=k}^{n} \binom{m}{k} = \binom{n+1}{k+1}$. *Hint: Note that summand on the right corresponds to the number of ways to select k elements out of $n + 1$; m in the formula on the left denote the maximum of this selected set minus one.*

11.10 Using the previous formula, find the formulas for the following expressions:

- $\sum_{k=0}^{n} k$,
- $\sum_{k=0}^{n} k^2$, and
- $\sum_{k=0}^{n} k^3$.

11.11 Show that

$$\sum_{k=0}^{n} \binom{2n}{2k} = \sum_{k=0}^{n-1} \binom{2n}{2k+1}$$

and

$$\sum_{k=0}^{n} \binom{2n+1}{2k} = \sum_{k=0}^{n} \binom{2n+1}{2k+1}$$

11.12 Show that $\sum_{k=0}^{n} \binom{m+k}{k} = \binom{m+n+1}{n}$.

11.13 Show that $\sum_{k=0}^{n} \binom{n-k}{k} = f_{n+1}$, where $f_1 = 1$, $f_2 = 1$, and $f_{n+2} = f_{n+1} + f_n$ for $n > 0$.

11.14 Show that $\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}$.

11.15 Show that $(a+1)^p \equiv a^p + 1 \pmod{p}$. *Hint: Use the binomial therem.*

11.16 We say that a function $f : \{0,1\}^n \to \{0,1\}$ depends on the ith argument iff for some $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in \{0,1\}$

$$f(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) \neq f(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n).$$

We also say that the function f depends on all the arguments iff for all $i \in [n]$ it depends on ith argument.

Find the number of functions $f : \{0,1\}^n \to \{0,1\}$ depending on all arguments.

11.17 Find the largest coefficient of $(x_1 + x_2 + \cdots + x_k)^k$.

11.18 Prove that, without using Theorem 11.6,

$$\sum_{k_1,k_2,k_3 \,:\, k_1+k_2+k_3=n} \binom{n}{k_1\ k_2\ k_3} = 3^n.$$

12 Partitions

The main question we study in this chapter is as follows: "how many ways to put n objects into k boxes". Note that there are four modes for this question:

1. the objects and boxes are identical,
2. the objects are identical but boxes are different,
3. the objects are different but boxes are identical,
4. the objects and boxes are different.

We are going to study the question in all these modes. The Table **??** summarizes the results we are going to prove for the cases when all the boxes are not empty.

12.1 Set Partitions

This section considers the case when objects are not identical.

First, we define a notion that allows us to compute the asnwer in case when all the boxes are the same.

Definition 12.1. *A partition of the set $[n]$ is a collection of non-empty blocks so that each element of $[n]$ belongs to exactly one of these blocks. The number of partitions of $[n]$ into k nonempty blocks is denoted by $S(n,k)$. The numbers $S(n,k)$ are called the* Stirling numbers of the second kind.

It is easy to see that $S(n,1) = 1$ and $S(n,n) = 1$. Moreover, $S(n,k) = 0$ if $k > n$ or $k \leq 0$.

Let us find the value in a more complicated setting, we claim that $S(n,n-2) = \binom{n}{2}$. Indeed, any partition of $[n]$ into $n-1$ blocks consists of $n-1$ singletons and one set with two elements, thus we just need to select these two elements.

Using double counting, one may prove a recursive formula for Stirling numbers of the second kind.

Theorem 12.1. *For any $n > k > 0$,*

$$S(n,k) = S(n-1,k-1) + S(n-1,k).$$

Object's name	Parameters	Formula
Surjections	n distinct objects k distinct boxes	$S(n,k)k!$
	n distinct objects any number of boxes	$\sum_{k=i}^{n} S(n,k)k!$
Compositions	n distinct objects k distinct boxes	$\binom{n-1}{k-1}$
	n distinct objects any number of boxes	2^{n-1}
Set partitions	n distinct objects k distinct boxes	$S(n,k)$
	n distinct objects any number of boxes	$B(n)$
Integer partitions	n distinct objects k distinct boxes	$p_k(n)$
	n distinct objects any number of boxes	$p(n)$

Table 12.1: Formulas for the cases when boxes are not empty

Note that the number of surjections from $[n]$ to $[k]$ is equal to the number of ways to put n different objects into k different boxes.

Lemma 12.1. *There are exactly $k!S(n,k)$ surjective functions from $[n]$ to $[k]$.*

Using this equality, we can prove a surprising result.

Theorem 12.2. *For any real x and positive integer n,*

$$x^n = \sum_{k=0}^{n} S(n,k)(x)_k,$$

where $(x)_k = \prod_{i=0}^{k-1}(x-i)$.

Definition 12.2. *The number of all set partitions of $[n]$ into nonempty parts is denoted by $B(n)$, and is called the nth* Bell number. *(We deifine $B(0) = 0$).*

It is easy to see that the following theorem holds.

Theorem 12.3. *For any $n \geq 0$,*

$$B(n+1) = \sum_{i=0}^{n} \binom{n}{i} B(i).$$

12.2 Compoisiton

This section answers the question in the case when the objects are the same but boxes are different. Since all the objects are identical, only the number of objects in each box matters.

Definition 12.3. *A sequence $(a_1, \ldots, a_k)$ of nonnegative integers such that $a_1 + \cdots + a_k = n$ is called a* weak composition *of n into k. If, in addition, all the numbers are positive, the sequence is called* composition*.*

Using the binomial coefficients we can find the number of weak compositions.

Theorem 12.4. *For all positive integers n and k, the number of weak compositions of n into k is equal to $\binom{n+k-1}{n}$.*

As a result, we can count the number of compositions.

Corollary 12.1. *For all positive integers n and k, the number of compositions of n into k is equal to $\binom{n}{n}$.*

Exercise 12.1. *Let $\ell_1, \ldots, \ell_k$ be some nonnegative numbers such that $\ell_1 + \cdots + \ell_k$. Find the number of weak compositions $(a_1, \ldots, a_k)$ of n into k such that $a_i \geq \ell_i$.*

Corollary 12.2. *The number of all compositions of n is equal to 2^{n-1}.*

12.3 Integer Partitions

Now consider the case when both objects and boxes are identical. In this case, as in the previous we are only intrested in the numbers of objects in boxes, but in addition, we are not intrested in the order of these numbers.

Definition 12.4. *Let n and $a_1 \geq a_2 \geq \cdots \geq a_k \geq 1$ be integers so that $a_1 + \ldots a_k = n$. Then the sequence $(a_1, \ldots, a_k)$ is called a* partition[1] *of the integer n.*

The number of all the partitions is denoted by $p(n)$ and the number of partitions of n into k parts is denoted by $p_k(n)$.

There is no good formula allowing to find the value of $p(n)$. However, we will prove some properties of $p(n)$. The main tool to explain proofs we are going to discuss are Young diagrams[2]. A Young diagram for a partition $(a_1, \ldots, a_k)$ consists of k columns of squares called "boxes" such that in the ith column there are a_i boxes (an example of such a diagram is depicted on 12.1). We can reflect a Young diagram of a partition of n with respect to its main diagonal, we get another shape, representing the *conjugate* partition of n (an example of such transformation is depicted on 12.1).

Using these diagrams it is easy to show the following theorems.

[1] Note that we used the word partition in two different meanings: one to denote a partition of a set $[n]$ and another to denote the partition of an integer n. In most of the cases the meaning is clear from the context; however, if it is necessary to emphasize that we mean partition of a set, we say set-partition. Note that in some languages there are two different words for these two notions; e.g in French "partition" is used for set-partitions, and "partage" for partitions of the integer n).

[2] A small variation of these diagrams is called Ferrers shapes after an american mathematician Norman Macleod Ferrers.

(a) The Young diagram for the partition $(4, 3, 1, 1)$.

(b) The conjugate of the Young diagram for the partition $(4, 3, 1, 1)$.

Figure 12.1: Young diagrams.

Theorem 12.5. *The number of partitions of n into at most k parts is equal to that of partitions of n into parts not larger than k.*

Theorem 12.6. *Let $q(n)$ be the number of partitions of n in which each part is at least two. Then $q(n) = p(n) - p(n-1)$, for all positive integers $n \geq 2$.*

Part III

Introduction to Mathematical Logic

13 Propositional Logic

In the Part I we studied the most important mathematical notation and how to prove theorems, we gave several "formal" definitions; however, our definition of the proof was not quite formal. It just allowed us to distinguish between good arguments and bad arguments. It is not enough if we wish to study proofs which is the core of mathematical logic.

Propositional Formulas:
Introduction to Mathematical Logic #1

youtu.be/X0797bVFf3Y

13.1 Propositional Formulas

First, we need to define what precisely we mean by "mathematical statement". In order to do this we define a class of formulas called propositional formulas.

Definition 13.1. *We say that ϕ is a propositional formula on the variables $x_1, \dots, x_n$ if*

- *either ϕ is equal to x_i for some $i \in [n]$,*
- *or ϕ is equal to $(\psi_1 \wedge \psi_2)$ or $(\psi_1 \vee \psi_2)$, where ψ_1 and ψ_2 are propositional formulas on $x_1, \dots x_n$,*
- *or ϕ is equal to $\neg\psi$, where ψ is a propositional formula on $x_1, \dots x_n$.*

We can also define the value of a formula.

Definition 13.2. *Let ϕ be a propositional formula on the variables $x_1, \dots,$ x_n and $v_1, \dots, v_n \in \{T, F\}$.*

We say that the value $\phi|_{x_1=v_1,\dots,x_n=v_n}$ of the formula ϕ when v_1 is substituted as the value of $x_1, \dots,$ and v_n is substituted as the value of x_n is equal to

- *v_i if ϕ is equal to x_i, and*
- *$\psi_1|_{x_1=v_1,\dots,x_n=v_n} \wedge \psi_2|_{x_1=v_1,\dots,x_n=v_n}$ if ϕ is equal to $(\psi_1 \wedge \psi_2)$, and*
- *$\psi_1|_{x_1=v_1,\dots,x_n=v_n} \vee \psi_2|_{x_1=v_1,\dots,x_n=v_n}$ if ϕ is equal to $(\psi_1 \vee \psi_2)$, and*
- *$\neg\psi|_{x_1=v_1,\dots,x_n=v_n}$ if ϕ is equal to $\neg\psi$.*

For example, the value of a formula $(x_1 \wedge x_2) \vee x_3$ when T is substituted as the value of x_1, T is substituted as the value of x_2, and F is substituted as the value of x_3 is equal to $(\mathrm{T} \wedge \mathrm{T}) \vee \mathrm{F} = \mathrm{T}$.

Theorem 13.1. *For any function $f : \{T, F\}^n \to \{T, F\}$ there is a formula ϕ on the variables $x_1, \ldots, x_n$ such that $\phi\big|_{x_1=v_1,\ldots,x_n=v_n} = f(v_1, \ldots, v_n)$ for all $v_1, \ldots, v_n \in \{T, F\}$*

Proofs Using Truth Tables:
Introduction to Mathematical Logic #2

youtu.be/DOAxnmScpPc

13.2 *Truth Tables*

Let us start the discussion of mathematical logic from an example similar to the proof we gave in the begining of the first chapter. Assume that we know that if x is a real number such that $x < -2$ or $x > 2$, then $x^2 > 4$. We can derive that if $\neg(x^2 > 4)$, then $\neg(x < -2)$ and $\neg(x > 2)$.

In order ot emphasize the logical structure of the argument let us denote $x > 2$ by p, $x < -2$ by q, and $x^2 > 4$ by r. In this case the argument is as follows. Assume that we know that $(p \vee q) \implies r$. We can derive that $\neg r \implies (\neg p \wedge \neg q)$.

How can we check that this argument is correct? The simplest way is to use a truth table to check that whether the assumption is true, the consequence is also true.

p	q	r	$(p \vee q) \implies r$	$\neg r \implies (\neg p \wedge \neg q)$
T	T	T	T	T
T	T	F	F	F
T	F	T	T	T
T	F	F	F	F
F	T	T	T	T
F	T	F	F	F
F	F	T	T	T
F	F	F	T	T

It is easy to note that the argument is indeed correct, i.e. if $(p \vee q) \implies r$ is true, then $\neg r \implies (\neg p \wedge \neg q)$ is also true. This statement says that

$$\big((p \vee q) \implies r\big) \iff \big(\neg r \implies (\neg p \wedge \neg q)\big)$$

is always true (we say that this propositional formula is a *tautology*). A generalization of this saying the if $p \implies q$ is true, then $\neg q \implies \neg p$ is also true is called the *contraposition* argument.

Let us now consider another argument. If we know that Joe was a good boy and we know that if Joe is a good boy, then Santa gives a present to Joe. We may conclude that Santa gives a present to Joe.

We can similarly to the previous example write this argument using variables and connectives. If we know that p and $p \implies q$, we may conclude that q is true.

Exercise 13.1. *Show that this argument is correct.*

Such an argument is called *modus ponens*.

Natural Deduction:
Introduction to Mathematical Logic #3

youtu.be/PfVafyptFtM

13.3 Natural Deduction

The problem of this method is that we need to consider **all** possible values of the variables. Let us now consider a more complicated example. Imagine that we know that $\neg q$, $p \implies q$. Using the contraposition argument and modus ponens we may derive $\neg p$. Indeed, by contraposition we may conclude that $\neg q \implies \neg p$ and modus ponens impies that $\neg p$ is true since $\neg q$ is true.

In other words, we can combine several tautologies to prove another tautology. Apparently it is enough to fix some small number of tautlogies to derive all other tautologies, we call these tautlogies "rules". There are several ways to write such proofs, we are going to use Fitch notation for natural deduction. In this notation any proof is writen in several rows, each row in a Fitch-style proof is either:

- an assumption or subproof assumption.
- a sentence justified by the citation of (1) a rule of inference and (2) the prior line or lines of the proof that license that rule.

We say that there is a natural deduction derivation of ϕ from $\psi_1, \ldots,$ ψ_k. If there is a Fitch-style proof starting with the assumptions $\psi_1, \ldots,$ ψ_k, and finishes with the formula ϕ. Using this scheme we may write the argument we just mentioned as follows.

1	$\neg q$	
2	$p \implies q$	
3	$\neg q \implies \neg p$	contraposition, 2
4	$\neg p$	modus ponens, 1, 3

In the rest of the section we are going to list all the rules we use.

Conjunctions. In order to introduce a conjunction we can use the following rule.

m	A	
n	B	
	$A \wedge B$	$\wedge$I, m, n

This rule corresponds to the tautology $(A \wedge B) \implies (A \wedge B)$.

In order to eliminate conjunctions we can use the following two rules.

$$
\begin{array}{r|ll}
m & A \wedge B & \\
 & A & \wedge\mathrm{E},\, m
\end{array}
\qquad\qquad
\begin{array}{r|ll}
m & A \wedge B & \\
 & B & \wedge\mathrm{E},\, m
\end{array}
$$

These rules correspond to the tautologies $(A \wedge B) \implies A$ and $(A \wedge B) \implies B$.

Disjuctions. In order to introduce a disjunction we can use the following two rules.

$$
\begin{array}{r|ll}
m & A & \\
 & A \vee B & \vee\mathrm{I},\, m
\end{array}
\qquad\qquad
\begin{array}{r|ll}
m & A & \\
 & B \vee A & \vee\mathrm{I},\, m
\end{array}
$$

These rules correspond to the tautologies $A \implies (A \vee B)$ and $A \implies (B \vee A)$.

In order to eliminate a disjunction we can use the following rule.

$$
\begin{array}{r|ll}
m & A \vee B & \\
i & \quad \begin{array}{|l} A \\ \hline \end{array} & \\
j & \quad \begin{array}{|l} C \end{array} & \\
k & \quad \begin{array}{|l} B \\ \hline \end{array} & \\
l & \quad \begin{array}{|l} C \end{array} & \\
 & C & \vee\mathrm{E},\, m,\, i\text{–}j,\, k\text{–}l
\end{array}
$$

This rule corresponds to the tautology $\big((A \vee B) \wedge (A \implies C) \wedge (B \implies C)\big) \implies C$.

Implications. In order to introduce an implication we can use the following two rules.

$$
\begin{array}{r|ll}
i & \quad \begin{array}{|l} A \\ \hline \end{array} & \\
j & \quad \begin{array}{|l} B \end{array} & \\
 & A \implies B & \Rightarrow\mathrm{I},\, i\text{–}j
\end{array}
$$

This rule corresponds to the tautology $(A \implies B) \implies (A \implies B)$.

In order to eliminate an implication we can use the following rule.

$$\begin{array}{r|ll} m & A \implies B & \\ n & A & \\ & B & \Rightarrow\mathrm{E},\, m,\, n \end{array}$$

This rule corresponds to the tautology $\big((A \implies B) \wedge A\big) \implies B$.

Negations. In order to introduce a negation we can use the following two rules ($\perp$ is a special symbol representing a false statement).

$$\begin{array}{r|ll} i & \quad \begin{array}{|l} A \\ \hline \end{array} & \\ j & \quad \begin{array}{|l} \perp \end{array} & \\ & \neg A & \neg\mathrm{I},\, i\text{–}j \end{array}$$

This rule corresponds to the tautology $\big(A \implies \perp\big) \implies \neg A$.

In order to eliminate a negation we can use the following rule.

$$\begin{array}{r|ll} m & A & \\ n & \neg A & \\ & \perp & \neg\mathrm{E},\, m,\, n \end{array}$$

This rule corresponds to the tautology $\big(A \wedge \neg A\big) \implies \perp$.

Truths and falsities. Additionally, we have the following two rules.

$$\begin{array}{r|ll} m & \perp & \\ & A & \perp\mathrm{E},\, m \end{array} \qquad \begin{array}{r|ll} i & \quad \begin{array}{|l} \neg A \\ \hline \end{array} & \\ j & \quad \begin{array}{|l} \perp \end{array} & \\ & A & \mathrm{IP},\, i,\, j \end{array}$$

An online tool to check natural deduction proofs

`proofs.openlogicproject.org/`

Exercise 13.2. *Check that all the tautologies we mentioned are indeed tautologies.*

13.4 *Examples of Derivations*

In this section we give several derivations using the rules we just introduced.

First, we prove that if we know that $A \implies \neg A$ we can derive that $\neg A$.

1	$A \implies \neg A$	
2	$\quad A$	
3	$\quad \neg A$	$\Rightarrow$E, 1, 2
4	$\quad \bot$	$\neg$E, 2, 3
5	$\neg A$	$\neg$I, 2–4

Another statement we are going to prove is that if $A \implies (A \wedge \neg A)$ is true, then $\neg A$ is also true.

1	$A \implies (A \wedge \neg A)$	
2	$\quad A$	
3	$\quad A \wedge \neg A$	$\Rightarrow$E, 1, 2
4	$\quad \neg A$	$\wedge$E, 3
5	$\quad \bot$	$\neg$E, 2, 4
6	$\neg A$	$\neg$I, 2–5

A bit more complicated is the proof of the law of excluded middle: $A \vee \neg A$.

1		
2	$\quad \neg(A \vee \neg A)$	
3	$\quad\quad A$	
4	$\quad\quad A \vee \neg A$	$\vee$I, 3
5	$\quad\quad \bot$	$\neg$E, 2, 4
6	$\quad \neg A$	$\neg$I, 3–5
7	$\quad A \vee \neg A$	$\vee$I, 6
8	$\quad \bot$	$\neg$E, 2, 8
9	$A \vee \neg A$	IP, 2–8

13.5 Soundness and Completeness

The most important properties of the natural deduction are the following two theorems.

Theorem 13.2 (completeness of natural deductions)**.** *Let ϕ be a propositional formula. If ϕ is a tautology, then there is a proof of ϕ.*

Theorem 13.3 (soundness of natural deductions)**.** *Let ϕ be a propositional formula. If there is a proof of ϕ, then ϕ is a tautology.*

Proofs of these two theorems are not that difficult but very technicall. So prove these statements on examples to at least illustate them.

Completeness of natural deductions. Proofs of this statement exploit the following idea: if a propositional formula is a tautlogy, then we can verify this statement using the truth table. So the proof simply brute-force all the values of the variables of a formula and check that the formula is indeed true. Consider a tautlogy $(\neg A \land \neg B) \implies \neg(A \lor B)$. The proof of this tautlogy is as follows.

First we derive $A \lor \neg A$ and $B \lor \neg B$, we use these two formulas to consider casess using the elimination of disjunction.

1		
2	$A \lor \neg A$	the law of excluded middle
3	$B \lor \neg B$	the law of excluded middle

After that we consider the case when A and B are both true. Note that the assumption of the implication is false in this case. Thus we just need to assume $\neg A \land \neg B$, derive the contradiction, and derive $\neg(A \lor B)$.

4	A	
5	B	
6	$\neg A \land \neg B$	
7	$\neg A$	$\land$E, 6
8	$\bot$	$\neg$E, 4, 7
9	$\neg(A \lor B)$	$\bot$E, 8
10	$(\neg A \land \neg B) \implies \neg(A \lor B)$	$\Rightarrow$I, 6–9

After that we consider the case when A is true but and B is false. In this case the assumption of the implication is also false, thus the proof is the same as in the previous case.

11	$\neg B$	
12	$\neg A \wedge \neg B$	
13	$\neg A$	$\wedge$E, 12
14	$\perp$	$\neg$E, 4, 13
15	$\neg(A \vee B)$	$\perp$E, 14
16	$(\neg A \wedge \neg B) \implies \neg(A \vee B)$	$\Rightarrow$I, 6–9
17	$(\neg A \wedge \neg B) \implies \neg(A \vee B)$	$\vee$E, 2, 5–10, 11–16

The third case is when A is false but and B is true. In this case the assumption of the implication is false again, thus the proof is the same as in the previous two cases.

18	$\neg A$	
19	B	
20	$\neg A \wedge \neg B$	
21	$\neg B$	$\wedge$E, 20
22	$\perp$	$\neg$E, 19, 22
23	$\neg(A \vee B)$	$\perp$E, 22
24	$(\neg A \wedge \neg B) \implies \neg(A \vee B)$	$\Rightarrow$I, 20–23

Finally, we consider the case when A and B are false. In this case the assumption of the implication is true, and since the formula is a tautlogy and $\neg A \wedge \neg B$ is true, we know that $\neg(A \vee B)$ is also true. Assume that $A \vee B$ is true and note that this is impossible. Thus using introudction of the negation we can prove the statement.

25	$\neg B$	
26	$\neg A \wedge \neg B$	
27	$A \vee B$	
28	A	
29	$\perp$	$\neg$E, 18, 28
30	B	
31	$\perp$	$\neg$E, 25, 30
32	$\perp$	$\vee$E, 27, 28–29, 30–31
33	$\neg(A \vee B)$	$\neg$E, 26–32
34	$(\neg A \wedge \neg B) \implies \neg(A \vee B)$	$\Rightarrow$I, 26–33
35	$(\neg A \wedge \neg B) \implies \neg(A \vee B)$	$\vee$E, 1, 3–17, 18–34

Soundness of natural deductions. Idea behind the soundness is also simple. We just explain that every line of the proof represent a tautology, including the last one. We illustrate this on the exaple of the proof of $A \vee \neg A$. Recall that the proof of this tautology is the following.

1	
2	$\neg(A \vee \neg A)$
3	A
4	$A \vee \neg A$
5	$\perp$
6	$\neg A$
7	$A \vee \neg A$
8	$\perp$
9	$A \vee \neg A$

1. The second line is just an assumption, so the corresponding tautology is $\neg(A \vee \neg A) \implies \neg(A \vee \neg A)$.

2. Line 3 is also an assumption so the corresponding tautology is $\neg(A \vee \neg A) \implies (A \implies A)$.

3. Line 4 is a formula $A \vee \neg A$ which we derived under assumptions $\neg(A \vee \neg A)$ and A, so the corresponding tautlogy is $\neg(A \vee \neg A) \implies$

$(A \implies (A \vee \neg A))$ (it is a tautology since we replaced A by $A \vee \neg A$ in the conclusion of the formula corresponding to Line 3).

4. Line 5 is a formula $\perp$ which we derived under assumptions $\neg(A \vee \neg A)$ and A, so the corresponding tautlogy is $\neg(A \vee \neg A) \implies (A \implies \perp)$ (it is a tautology since on Line 4 we explained that $\neg(A \vee \neg A) \implies (A \implies (A \vee \neg A)))$.

5. Line 6 is a formula $\neg A$ which we derived under assumptions $\neg(A \vee \neg A)$, so the corresponding tautlogy is $\neg(A \vee \neg A) \implies \neg A$ (it is a tautology since on Line 5 we explained that $A \implies \perp$ under the assumption $\neg(A \vee \neg A)$).

6. Line 7 is a formula $A \vee \neg A$ which we derived under assumptions $\neg(A \vee \neg A)$, so the corresponding tautlogy is $\neg(A \vee \neg A) \implies (A \vee \neg A)$ (it is a tautology since on Line 6 we explained that A under the assumption $\neg(A \vee \neg A)$).

7. Line 8 is a formula $\perp$ which we derived under assumptions $\neg(A \vee \neg A)$, so the corresponding tautlogy is $\neg(A \vee \neg A) \implies \perp$ (it is a tautology since on Line 6 we explained that $A \vee \neg A$ under the assumption $\neg(A \vee \neg A)$).

8. Finally, Line 9 is a formula $A \vee \neg A$ (it is a tautlogy since we proved that $\neg(A \vee \neg A) \implies \perp$ is a tautlogy)

End of The Chapter Exercises

13.3 Derive $(A \wedge B) \implies C$ from $A \implies (B \implies C)$.

13.4 Derive $A \vee C$ from $(A \wedge B) \vee C$.

13.5 Derive $B \vee C$ from $A \implies B$ and $\neg A \implies C$.

14 Predicate Logic

In the previous chapter we defined natural deductions for propositional logic. But in real mathematics there are many formulas that are not propositional. For example we may wish to prove that if a relation R on M is transitive, then

$$(R(w,x) \wedge R(x,y) \wedge R(y,z)) \implies R(w,z)$$

is true for any $w, x, y, z \in M$. In this chapter we define a logical system that allows us to formally prove such statements.

14.1 Predicate Formulas

Let us formalize the previous statement:

$$\begin{aligned}&(\forall x, y, z \in M\ (R(x,y) \wedge R(y,z)) \implies R(x,z)) \implies \\ &\qquad (\forall w, x, y, z \in M\ (R(w,x) \wedge R(x,y) \wedge R(y,z)) \implies R(x,z))\end{aligned}$$

Note that there are several things we need to explain if we wish to define formally formulas like this:

- we need to explain what kind of sets we can use (in this case we need to define M),
- we need to explain what kind of relations we can use (in this case we need to define R),

Another example of a formula we may wish to prove is saying that if $f : M \to M$ is an inverse of itself (i.e. $f(f(x)) = f(x)$ for any $x \in M$), then $f(f(f(x))) = f(x)$ for any $x \in M$; more formaly, we may wish to prove a statement

$$(\forall x \in M\ f(f(x)) = x) \implies (\forall x \in M\ f(f(f(x))) = f(x)).$$

In order to expalin what we mean by such formulas

- we need to explain what kind functions we can use (in this case we need to define f).

Signature. In predicate logic, formula uses just symbols for all these objects. We specify these symbols only when we wish to compute actual truth value of the formula. We also assume that all the quantifiers are over the same set so we do not need a symbol for the set M.

Signature is the way to define the list of all these symbols, it consists of three objects:

- the set of symbols for relations,
- the set of symbols for functions,
- arities of these functions and relations (i.e. how many arguments they may take).

An example of a signature is a triple $(\{\text{“R”}\}, \{\text{“f”}\}, \mathrm{ar})$, where

$$\mathrm{ar}(s) = \begin{cases} 2 & \text{if } s = \text{“R”} \\ 1 & \text{if } s = \text{“f”} \end{cases}.$$

This signature is enough to define the formulas we discussed. Now we are ready to defin the predicate formulas.

Definition 14.1. *Let $\mathcal{S} = (S_{\mathrm{rel}}, S_{\mathrm{fun}}, a)$ be a signature.*

We say that t is a term in the signature $\mathcal{S}$ over the variables $x_1, \ldots, x_n$ if

- *either t is equal to a variable x_i*
- *or t is equal to $f(t_1, \ldots, t_\ell)$, where $f \in S_{\mathrm{fun}}$, $\ell = a(f)$, and $t_1, \ldots, t_\ell$ are terms in the signature $\mathcal{S}$.*

We say that ϕ is a predicate formula in the signature $\mathcal{S}$ over the variables $x_1, \ldots, x_n$ if

- *either ϕ is equal to $R(t_1, \ldots, t_\ell)$, where $R \in S_{\mathrm{rel}}$, $\ell = a(R)$, and $t_1, \ldots, t_\ell$ are terms in the signature $\mathcal{S}$.*
- *or ϕ is equal to $(\psi_1 \wedge \psi_2)$ or $(\psi_1 \vee \psi_2)$, where ψ_1 and ψ_2 are predicate formulas in the signature $\mathcal{S}$,*
- *or ϕ is equal to $\neg\psi$, where ψ is a predicate formula in the signature $\mathcal{S}$,*
- *or ϕ is equal to $\exists x_i\ \psi$ or $\forall x_i\ \psi$ where ψ is a predicate formula in the signature $\mathcal{S}$.*

In order to compute the truth value of a predicate formula, we need to specify the values of all the free variables and all the symbols from the signature. The specification of the symbols from the signature is called structure; i.e. a structure for a signature $\mathcal{S} = (S_{\mathrm{rel}}, S_{\mathrm{fun}}, a)$ is a triple $(M, F_{\mathrm{rel}}, F_{\mathrm{fun}})$ such that

- $F_{\mathrm{rel}} : S_{\mathrm{rel}} \to \bigcup_{i=0}^{\infty} 2^{M^i}$ such that $F_{\mathrm{rel}}(R) \in 2^{M^{a(R)}}$ and

- $F_{\text{fun}} : S_{\text{fun}} \to \bigcup_{i=0}^{\infty} M^{M^i}$ such that $F_{\text{fun}}(f) \in M^{M^{a(f)}}$.

The set M in the structure is called the domain of the structure.

Definition 14.2. *Let $\mathcal{S} = (S_{\text{rel}}, S_{\text{fun}}, a)$ be a signature and $\mathcal{M} = (M, F_{\text{rel}}, F_{\text{fun}})$ be a structure for $\mathcal{S}$.*

Let t be a term in the signature $\mathcal{S}$ over the variables $x_1, \dots, x_n$ and $v_1, \dots, v_n \in M$. The value of t with $x_1 = v_1, \dots, x_n = t_n$ with respect to the structure $\mathcal{M}$ is equal

- *either to v_i when $t = x_i$,*

- *or $F_{\text{fun}}(f)(\mu_1, \dots, \mu_{a(f)})$ when $t = f(t_1, \dots, t_{a(f)})$, where μ_i is equal to the value of t_i with $x_1 = v_1, \dots, x_n = v_n$ with respect to the structure $\mathcal{M}$.*

Let ϕ be a formula in the signature $\mathcal{S}$ over the variables $x_1, \dots, x_n$.

- *Let ϕ be equal to $F_{\text{rel}}(R)(t_1, \dots, t_{a(R)})$, where $t_1, \dots, t_n$ are some terms in $\mathcal{S}$. Then the value of ϕ with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$ is equal to $R(\mu_1, \dots, \mu_{a(R)})$, where μ_i is equal to the value of t_i with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$.*

- *Let ϕ be equal to $\psi_1 \# \psi_2$, where $\# \in \{\vee, \wedge\}$ and ψ_1, ψ_2 are predicate formulas. Then the value of ϕ with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$ is equal to $\beta_1 \# \beta_2$, where β_i is equal to the value of ψ_i with $x_1 = v_1$, $\dots, x_n = v_n$ with respect to $\mathcal{M}$.*

- *Let ϕ be equal to $\neg\psi$, where ψ is a predicate formula. Then the value of ϕ with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$ is equal to $\neg\beta$, where β is equal to the value of ψ with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$.*

- *Let ϕ be equal to $\exists x_i\ \psi$, where ψ is a predicate formula. Then the value of ϕ with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$ is equal to true iff there is $\mu \in M$ such that the value of ψ with $x_1 = v_1, \dots, x_{i-1} = v_{i-1}$, $x_i = \mu$, $x_{i+1} = v_{i+1}, \dots x_n = v_n$ with respect to $\mathcal{M}$.*

- *Let ϕ be equal to $\forall x_i\ \psi$, where ψ is a predicate formula. Then the value of ϕ with $x_1 = v_1, \dots, x_n = v_n$ with respect to $\mathcal{M}$ is equal to true iff for all $\mu \in M$, the value of ψ with $x_1 = v_1, \dots, x_{i-1} = v_{i-1}$, $x_i = \mu$, $x_{i+1} = v_{i+1}, \dots x_n = v_n$ with respect to $\mathcal{M}$.*

Let us consider an example:

- First, we define a signature $\mathcal{S} = (\{=, <\}, \{+, \cdot\}, \text{ar})$, where $\text{ar}(x) = 2$ for any $x \in \{=, +, \cdot\}$.

- After this we define a structure $\mathcal{M} = (\mathbb{R}, F_{\text{rel}}, F_{\text{fun}})$, where

$$F_{\text{fun}}(f)(x,y) = \begin{cases} x \cdot y & \text{if } f \text{ is } \cdot \\ x + y & \text{if } f \text{ is } + \end{cases}$$

and

$$F_{\text{rel}}(R)(x,y) \begin{cases} x = y & \text{if } R \text{ is } = \\ x < y & \text{if } R \text{ is } < \end{cases}$$

- Finally, we consider the formulas in the signature $\mathcal{S}$

$$\forall x\ \forall y\ x + y = y + x$$

and

$$\forall x\ \forall y\ \forall z\ (x < y \implies x + z < y + z).$$

(Note that we write $a = b$ instead of $=(a, b)$ and $a + b$ instead of $a +$ b, this is a common notation when we the standard mathematical operations and relations are used in the signature.)

The first formula says that the operation $+$ is commutative, which is true, so the value of the formula with respect to the structure $\mathcal{M}$ should be true. (Note that we do not mention the values of the variables x and y since both of them are not free.) Indeed, consider $a, b \in \mathbb{R}$ note that the value of $x + y = y + x$ with $x = a$ and $y = b$ and with respect to the structure $\mathcal{M}$ is equal to $F_{\text{rel}}(F_{\text{fun}}(a, b), F_{\text{fun}}(b, a))$ which is the same as $a + b = b + a$; thus, the first formula is true.

The second formula says that the inequalities are additive, so it should be also true with respect to the structure $\mathcal{M}$.

Exercise 14.1. *Show that the second formula is true with respect to the structure $\mathcal{M}$.*

By analogy with the tautology, in the predicate logic we wish to prove that a formula is true, whenever the structure and the values of the variables we choose. Such formulas a called *logically valid*.

14.2 Natural Deduction

Natural deduction for the predicate formulas is defined in the same maner as the natural deduction for the propositional formulas but now the lines are predicate formulas and we can use four additional rules.

Universal quantifier. The first logically-valid formula we use as a rule is $A(x) \implies (\forall y\ A(y))$, this rule allows us to introduce an universal quantifier. In order to use the following rule, x should not be a free variable of an open hypothesis.

$$\begin{array}{r|ll}
m & A(x) & \\
 & \forall y\ A(y) & \qquad \forall\text{I}, m
\end{array}$$

The second logically-valid formula we use as a rule says that if a statement is true for all the values of a variable, then it is also true when you substitute some specific term instead of the variable, i.e. $(\forall x\ A(x)) \implies A(t))$, this rule allows us to eliminate an universal quantifier.

$$\begin{array}{r|ll}
m & \forall x\ A(x) & \\
 & A(t) & \qquad \forall\text{E}, m
\end{array}$$

Existential quantifier. The first formula for the exisential quantifier says that you can name any term in the formula by a variable and formula is still true for some value of the variable. The corresponding formula is $A(t) \implies (\exists x\ A(x))$.

$$\begin{array}{r|ll}
m & A(t) & \\
 & \exists x\ A(x) & \qquad \exists\text{I}, m
\end{array}$$

The last rule says that if $A(x)$ is true for some x and we know that $A(y)$ implies B, then we can derive B (note that this is true only when y is not used in B). Thus we can apply the following rule when y is not be a free variable neither of B nor of any open hypothesis.

$$\begin{array}{r|ll}
m & \exists x\ A(x) & \\
i & \quad \vert\ \underline{A(y)} & \\
j & \quad \vert\ B & \\
 & B & \qquad \exists\text{E}, m, i\text{–}j
\end{array}$$

14.3 Examples of Derivations

First example $\forall x\ F(x) \vee \neg(\forall x\ F(x))$ is a special form of the law of excluded middle, which we proved in the previous chapter. However, in order to emphasize that the propositional logic can prove all the statements provable in the predicate case we present the proof of this statement as well.

1		
2	$\quad \neg(\forall x\, F(x) \vee \neg(\forall x\, F(x)))$	
3	$\quad\quad \forall x\, F(x)$	
4	$\quad\quad \forall x\, F(x) \vee \neg(\forall x\, F(x))$	$\vee$I, 3
5	$\quad\quad \bot$	$\neg$E, 2, 4
6	$\quad \neg(\forall x\, F(x))$	$\neg$I, 3–5
7	$\quad \forall x\, F(x) \vee \neg(\forall x\, F(x))$	$\vee$I, 6
8	$\quad \bot$	$\neg$E, 2, 8
9	$\forall x\, F(x) \vee \neg(\forall x\, F(x))$	IP, 2–8

Unfortunately, this example just shows that a statement provable in the propositional logic can be proven in the predicate logic. The next example is an example that cannot be expressed in the propositional logic, we prove that if we know that $\forall x \forall y\, R(x,y) \implies R(y,x)$, the we can derive $\forall x \forall y\, ((R(x,y) \implies R(y,x)) \wedge (R(y,x) \implies R(x,y)))$.

1	$\forall x \forall y\, R(x,y) \implies R(y,x)$	
2	$\forall y\, R(x',y) \implies R(y,x')$	$\forall$E, 1
3	$R(x',y') \implies R(y',x')$	$\forall$E, 2
4	$\forall y\, R(y',y) \implies R(y,y')$	$\forall$E, 1
5	$R(y',x') \implies R(x',y')$	$\forall$E, 4
6	$(R(x',y') \implies R(y',x')) \wedge R(y',x') \implies R(x',y')$	$\wedge$I, 3, 5
7	$\forall y\, (R(x',y) \implies R(y,x')) \wedge (R(y,x') \implies R(x',y))$	$\forall$I, 7
8	$\forall x \forall y\, (R(x,y) \implies R(y,x)) \wedge (R(y,x) \implies R(x,y))$	$\forall$I, 7

14.4 Soundness and Completeness

Like in the propositional case, the most important properties of the natural deduction are the following two theorems.

Theorem 14.1 (completeness of natural deductions)**.** *Let ϕ be a predicate formula. If ϕ is logically valid, then there is a proof of ϕ.*

Theorem 14.2 (soundness of natural deductions)**.** *Let ϕ be a predicate formula. If there is a proof of ϕ, then ϕ is logically valid.*